WALTON, ELIZABETH M. Ph.D. Ecological Niche Modeling as a Conservation Tool to Predict Actual and Potential Habitat for the Bog Turtle, *Glyptemys muhlenbergii.* (2009)
Directed by Dr. Roy S. Stine and Dr. Rick L. Bunch. 147 pp.

The bog turtle (*Glyptemys muhlenbergii*) is faced with two principle threats: wetland habitat loss and, to a lesser degree, the illegal collection for pet trade demands. Current methodologies for bog turtle population discovery in the Southeast rely primarily on field surveys, which are labor intensive and fiscally exhaustive.

The purpose of this research was to evaluate the role of geographic information science technologies, remote sensing and ecological niche modeling to predict potential bog turtle habitats in the Southeast. Environmental data were organized in a geographic information system. The Genetic Algorithm for Rule-set Production was used to develop an ecological niche model to identify additional habitat sites with the same signatures and potential capacity for support.

The results showed the area under the curve as 97%; the model correctly predicted 98.889% of the data points; and the model predicted 1.67% of the total research area as potential habitat. Areas of highest prediction will be investigated for bog turtle occupancy by trained professionals. This information will be beneficial to researchers in setting conservation priorities for the bog turtle.

Keywords: ecological niche models, bog turtles, genetic algorithm, GARP, habitat prediction, geographic information systems, remote sensing, wetlands.

ECOLOGICAL NICHE MODELING AS A CONSERVATION TOOL TO

PREDICT ACTUAL AND POTENTIAL HABITAT FOR THE

BOG TURTLE, *GLYPTEMYS MUHLENBERGII*

by

Elizabeth M. Walton

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2009

Approved by

Dr. Roy S. Stine
Committee Co-Chair

Dr. Rick L. Bunch
Committee Co-Chair

This dissertation is dedicated to both my families,

The Morells and the Rosens

And especially to my parents,

Nell and Ray Morell

And in loving memory of my mother-in-law and father-in-law,

Molly and Jack Rosen

~~~~~~~~~~

To my son,

I miss your laughter and the warmth of your spirit
That would shine so brightly through your eyes; your stories
And pranks, your warm heart, the love and passions you had for life.

May you find the peace, fulfillment and happiness you seek,
And if you should ever venture beyond yourself,
I hope someday you'll find your way back to all of us.
~~~~~~~~~~

APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Co-Chair ______________________
Roy S. Stine

Committee Co-Chair ______________________
Rick L. Bunch

Committee Members ______________________
P. Daniel Royall

____
George R. Cline

____
Date of Acceptance by Committee

____
Date of Final Oral Examination

ACKNOWLEDGMENTS

This dissertation could not have been accomplished without the love, friendship, mentoring and support of my family and friends.

I will always remember the enthusiasm expressed by Mary Hall-Brown when she introduced me to the concepts of spatial analysis and remote sensing. She spent countless hours explaining the details and soaring on sheer enthusiasm for the science behind these technologies. Dr. L. Joe Morgan shared this passion and reinforced the potential of applying these technologies to bog turtles after giving me the article written by Dr. Chris Raxworthy (2003) describing his use of GARP in ecological niche modeling. If not for you two, I'm not sure I would have ever imagined the possibilities and I am forever grateful for your continued friendship.

Dr. Roy S. Stine has been my friend, mentor and one of my strongest supporters in the Geography Department at the University of North Carolina Greensboro. Roy, I owe you a great dept of gratitude. Not only did you help me over and around many obstacles found on this path, but you never lost faith in me. Even when it looked as though we wouldn't be able to conduct this research with the available data and hardware, you gave me the necessary tools (including a 64-bit super computer) and the freedom to explore the potential of this science while using it in a new context. You gave me license to try again and again until I was able to find my way through the applications and to develop the

methods required to use them successfully.

Dr. Rick L. Bunch agreed to be a co-chair on the dissertation committee and was the first to let me test the openModeller program on his “super” computer. Our chats and conversations led me to pursue this research with even more certainty that, somehow, it really could work. I appreciate all the times you have reviewed my work and given me positive feedback; it has only made this dissertation that much stronger.

Dr. Dan Royall, you have been my voice of encouragement, wisdom and reason. I remember when you told me I should stick to something I knew and was passionate about. I am so glad that I heeded your advice for not only did it benefit me, it is these words of wisdom that I now pass on to my own students. When I was so determined to follow my preconceived designs of creating a “perfect” model using all 1-meter data for this research, you taught me about the “less than perfect science,” and that sometimes we have to sacrifice the “ideal” model for a method that will work and still provide meaningful results. You also provided one of the most amazing learning experiences during my dissertation proposal defense when you helped me to realize how much personal passion I had injected into this research without holding first to its statistical significance.

Dr. George Cline, I am so fortunate to have you as my biologist on this dissertation committee. I appreciate your input and insight into the biology and ecology of this species as well as the scientific concepts behind environmental niches. Your questions and insight have challenged me to always remain faithful

to the biological concepts of bog turtles and ecological niche modeling while gaining deeper understanding and appreciation of the technologies used to simulate real-world conditions.

Dennis Herman has been a champion and passionate voice for bog turtles for many, many years. I found I could learn more from him about bog turtles in one day out in the field than I could by reading countless articles. I thank you, not only for all your tireless efforts on behalf of this species and Project Bog Turtle, but also for being so generous with your time and knowledge. John Behler (1997) wrote:

> *Despite problems [conserving species], there appear to be success stories . . . These have been long, arduous affairs, and for the most part, they have been the work of extraordinarily dedicated individuals, not legions of bureaucrats. I suspect that the turtle wars will be fought and won and lost by individual 'turtle men' and 'turtle women' who are on divine missions from the chelonian gods to save their species.*

You are one of the people John was talking about, an "extraordinarily dedicated individual," and the epitome of a "turtle man."

To David Moore, Kathy Wolfe, Judy Scurry and Al Rosen: I thank each of you for your incredible assistance in the field. David was instrumental in helping me with the collection of data points to create the wetland polygons; Kathy and Judy conducted plant inventories at the wetland sites. Judy, you are such a treasured friend and I can never thank you enough for sharing your home in the mountains with me on my trips to Ashe County. It is an area that simply restores the soul, and I will miss our long talks on the porch watching the New River, in all its grace, flowing to the sea. Al helped me to investigate the ground reference data points on several trips to Ashe County and I could not have collected all the data necessary without him there to help with all the logistics.

I am forever grateful for all the technical assistance provided by James Nelson, the lab director in the Geography Department at the University of North Carolina Greensboro. When I needed a dedicated computer to run these models, he gave me one. When I said I needed something with greater capacity and speed, he upgraded me to a computer he configured just for my purposes. When I said I needed *more* capacity and *more* power, he configured a new 64-bit machine and set me up once again. When this research started rolling and I saw how much storage space I was going to need, he made room for it on the departmental server. Every researcher should be blessed with this type of technical support, the enthusiasm and “can do” attitude that Jim has always provided. So to Jim I say, “Go forth and cringe no more for I will no longer darken your doorstep with these outrageous requests!”

Technical support was also provided by Tim Sutton from the Reference Center of Environmental Information (CRIA) in Campinas, Brazil. It was through his guidance and tutoring on implementing the GARP algorithm in openModeller that I discovered that I would “need a bigger boat” to get the openModeller interface to work properly. Both he and Renato De Giovanni worked patiently

with me as we implemented openModeller on the 64-bit computer. Trouble-shooting of openModeller and model implementation on an individual basis has always been handled with patience and the generous spirit of these two programmers who are devoted to the ecological niche modeling cause. Renato was my resource for the history of how the GARP algorithm has been implemented over the years. Kristina McNyset provided insight into model parameters and how to interpret model results. I am truly grateful for each of you and the contributions you've made to the success of this research. Any errors or misinterpretations in this document are exclusively my own.

To my friends and family members who frequently bolstered my waning spirit when the nights were too long and the days too short, you have truly defined the word "friendship." This is especially true of Dr. L. Joe Morgan, Debbie Shoffner and Dr. Tresa M. Saxton.

And finally, but certainly not least of all, I will forever be grateful to my husband, Al. Unless you have done it yourself, there are few people who can truly understand how challenging it can be to return to college as an adult to pursue three consecutive degrees. There have been moments that were absolutely fabulous and some that were . . . not so fabulous. Through it all, however, Al has always maintained his faith in my ability to fulfill this dream. Even when I thought I didn't have the stamina to possibly make it through yet another year, or another semester, he gave me the strength and courage to continue. He has celebrated the highs and lifted my spirits and resolve during the low times. Through it all, he has been my best friend and has made great sacrifices to ensure that this dream came true. His love and support have known no bounds and he has given me the necessary free-will, autonomy, and independence to pursue my aspirations regardless of where they may lead.

The Verification of Vulnerability: Bog Turtle

Guarded by horned beak and nails, surrounded
By mahogany carapace molded in tiles
Like beveled wood, hidden within the hingeless
Plastron, beneath twelve, yellow-splotched
Black scutes, buried below the inner lungs
And breast, harbored in the far reaches
Of the living heart, there it exists,
As it must, that particle of vulnerability,
As definite in its place as if it were a brief glint
Of steel, buried inside the body of the bog turtle.

And it is carried in that body daily, like a pinpoint
Of diamond in a dark pouch, through marshy fields
And sunlit seepages, and it is borne in that body,
Like a crystal of salt-light locked in a case
Of night, borne through snail-ridden reeds and pungent
Cow pastures in spring. It is cushioned and bound
By folds of velvet, by flesh and the muscle
Of dreams, during sleep on a weedy tussock all afternoon.
It is divided and bequeathed again in June, protected
By thick sap, by yolk meal and forage inside its egg
Encompassed by the walls of shell and nest.

Maybe I can imagine the sole intention present
In the steady movement of turtle breath filled
With the odor of worms this morning, stirring
Clover moisture at the roots. Maybe I can understand
How the body has taken form solely
Around the possibility of its own death,
How the entire body of the bog turtle
Cherishes and maintains and verifies the existence
Of its own crucial point of vulnerability exactly
As if that point were the only distinct,
Dimensionless instant of eternity ever realized.
And maybe I can guess what it is we own,
If, in fact, it is true: the proof of possession
Is the possibility of loss.

— Pattiann Rogers

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Introduction

This dissertation was an exploration of the role of geographic information science (GISc), remote sensing, and ecological niche modeling in predicting potential habitat sites for the bog turtle, *Glyptemys muhlenbergii.* The northern bog turtle population was listed as “threatened” under the Endangered Species Act (ESA) (1973) in 1997 citing loss of habitat as the primary cause of decline (United States Fish and Wildlife Service [USFWS] 1997). The southern population, however, was listed as “threatened due to similarity of appearance” as there was a paucity of baseline population data, and the overall population status and geographic distribution was unknown. Yet, while the southern population did not have sufficient data to warrant the same level of protection as the northern population, it was determined to be too difficult for officials to distinguish between them when enforcing the law since the two populations share many phenotypic traits. Since the 1997 ruling, researchers have been conducting ground surveys throughout the southern range in an attempt to locate additional bog turtle populations to determine the overall population status and to implement conservation measures and management strategies, where possible, to enhance the future viability of the southern species.

Bog turtles are physically small (3 – 3.75 inches) and reside in small, noncontiguous mountain wetlands that are heterogeneous, due in part to different stages of succession and farming practices (Walton 2002; Herman 2003). Ground surveys to discover potential habitats, however, are labor intensive, financially expensive (Vogiatzakis 2003; Rushton et al. 2004), and in cases where there are reduced population numbers, the rates of discovery success are marginal. Therefore, there was a need to develop a new methodology that would streamline the search and discovery process and to enable researchers to explore areas that have the greatest likelihood of providing all of the environmental factors required to support bog turtles.

In recent years, researchers have discovered the application of GISc, remote sensing and ecological niche modelinging technologies to model species habitats and population distributions (Davis et al. 1990; Aspinal 1995; Akçakaya 1996; Scott et al. 1996; Chen and Peterson 2000; Peterson et al. 2002; Venkataraman et al. 2002; Turner et al. 2003). Using these new technologies, for example, Raxworthy et al. (2003) created an ecological niche model using a geographic information system (GIS) that incorporated environmental characteristics with records of occurrence (both current and historical) to predict the distribution of rare chameleons in Madagascar. Upon investigating three areas with the highest predictive success rates, *seven new species* of chameleons were identified (Raxworthy 2003; Ramanujan 2004; NASA website).

A GIS is used to assemble environmental data proxies that represent vegetation indices, soil characteristics, temperature, aspect, and other variables that may directly or indirectly provide the requisite components to support the species of interest. Data layers can then be manipulated so that they are in the same projection, trimmed to the same geographic extent and then rasterized to the same spatial resolution. Species occurrence data points are added to signify where environmental conditions are sufficient to currently support bog turtle populations. This compilation of environmental variables and occurrence data points represents real-world conditions at the species' physical address, referred to as the species' *fundamental niche* by Hutchinson (1957), and represents a *theoretical* space in a GIS (Pearson 2007).

Ecological niche models use algorithms to analyze a theoretical environmental space consisting of these environmental data layers and a set of species occurrence data points. The data layers are stacked so that each pixel is aligned with its corresponding planimetric (*x, y)* position in each layer. The data points can then be associated with specific positions within the environmental data layers. Each data point will have specific environmental variable values, such as *elevation = 1000, sand = 33.3; clay = 33.3 and silt = 33.4* for each data layer. Thus, the identification of environmental tolerance ranges can be obtained by combining all values for each variable to determine the extremes within which the species currently exists. The ecological niche model can then analyze this theoretical environment to identify other areas within these stacked layers that fall within these established tolerance ranges. Those areas with highest levels of environmental similarity are then projected onto *geographic* space (Pearson 2007) in the real world where they can be investigated for habitat suitability as well as the presence or absence of bog turtles.

A number of studies have used ecological niche models to theoretically predict species occurrences (Raxworthy et al. 2003; Engler et al. 2004; McNyset 2005), species declines (Bond et al. 2006), suitable habitats and the spread of disease (Peterson et al. 2002; Peterson et al. 2004; Adjemian et al. 2006; Blackburn et al. 2007) and invasive species (Kluza and McNyset 2004). However, these studies address a global or regional scale using data resolutions that range from 1000 acres (Bond 2006) to 1 km data (Raxworthy 2003; McNyset 2005; Blackburn 2007). This dissertation represents the first known application an ecological niche model to not only predict the distribution of suitable habitat for a turtle species, but it is also the first time the genetic algorithm for rule-set production (GARP) (Stockwell and Peters 1999; Stockwell and Peterson 2002) has been used with high spatial resolution data ranging from 1 to 10 meters.

Chapter 2 presents the biology and ecology of the bog turtle as well as its geographic distribution, habitat preferences and a description of Ashe County, the focal area of this research. Chapter 2 also identifies the causes of decline, describes the central issues involved with protecting endangered species and includes a broader look at the world-wide decline of chelonian species.

GARP is used in this research and it is described in Chapter 3 along with a short review of previous work conducted using this algorithm. Chapter 3 identifies the conceptual evolution of the term “niche” and the theoretical foundation for niche modeling. Chapter 3 also contains the specific goals and hypotheses of this research.

Chapter 4 of this dissertation is an account of the methods and parameters used to conduct this research including a description of all the data layers, their source, data manipulation and processing. Descriptive statistics and omission/commission are also described in this Chapter.

Chapter 5 presents the results and interpretation of these research findings including those variables that were more successful in model accuracy and those data variables that were removed from the analysis. A total of 50 models were run using the GARP algorithm and the top 20 of these models were analyzed in a GIS. Descriptive statistics were used to identify the bog turtle’s environmental envelope. It is important to note, however, that due to the confidentiality surrounding endangered species, all specific localities and findings of this research have been omitted from this paper dissertation as they will become publicly available upon publication. Thus, all specific information will only be shared with authorized agencies and members of this dissertation committee.

To conclude this dissertation, Chapter 6 contains the conclusions, a discussion of the research findings, future directions and potential studies that could grow out of this current research. It is important to note that this dissertation is limited to one small spatial extent out of the southern geographic range of the bog turtle. There are still a number of counties that will need the same level of investigation in order to obtain an over-arching theory that defines the environmental envelope for this species and identifies where suitable bog turtle habitat can be found. This information will focus search efforts in a more efficient and effective manner.

1.2 Expected Significance and Broader Implications

The goals of this research were threefold: First, it analyzed the role of GISc, remote sensing and the GARP ecological niche model in detecting small, discontiguous wetland areas that may support bog turtles. Some of these wetlands are grazed by livestock, some are not. Other areas have been drained and ditched at some point in the past, and have since filled in again. Some of the wetlands are in different stages of succession. Because these wetlands lack homogeneity and are typically very small in size, traditional land classification schemes and remote sensing software programs alone cannot detect these areas (Walton 2006) across the landscape. Ecological niche models, however, have the ability to analyze a number of environmental proxies, such as vegetation, soil composition, and aspect, to detect other areas of similar composition and thus potential capacity for the support of bog turtle populations. The results of ecological niche models can be further analyzed in a GIS in such a manner as to scale down areas to those of highest predictions and thus focus search efforts in those areas. Understanding how these technologies can be best utilized for conservation strategies will be important to bog turtle preservation as well as the management strategies for other turtle species.

Secondly, this research was designed to identify the environmental tolerance ranges that currently support bog turtle populations within the study area. This will assist researchers in understanding the environmental means and extremes in which this species can exist; it will inform them of where they are most likely to find areas of suitable habitat; and will assist in management strategies. Lastly, this research proposed to identify where bog turtle habitats are *not* likely to be found due to a lack of suitable environmental conditions to support their biology and ecology.

It was anticipated that these goals would direct research efforts to those areas that have a greater likelihood of meeting environmental conditions and thus potential capacity for the support of bog turtles. Alternatively, it would eliminate a large range of areas where bog turtles are not likely to exist. The broader implications, however, could prove to be far-reaching and instrumental in turtle conservation, not only here in the United States, but in other areas of the world. There is currently a world-wide decline of turtle species as a result of human consumption, habitat loss, alteration or degradation, and pet trade demands (Behler 1997).

Used as a preferred food source and in traditional medicines in Asia, the trade in turtles is measured at more than 10—12 million individual turtles per year (Behler 1997; Salzberg 1998; Rhodin 2000) creating an urgent call to locate and protect remaining turtle species.

The Asian demand for turtles has already been felt in the United States as well (Behler 1997). North Carolina experienced a dramatic loss when collection went from 460 individual turtles in 2001 to more than 23,000 turtles (mainly aquatic species) in 2002 (North Carolina G.S. 113-333). The turtles taken from North Carolina were probably used to satisfy the demands of European pet trade markets or the food markets in Asian countries (Herman pers. comm., 11-Aug-05; Altherr and Freyer 2000). This "Asian Turtle Crisis," as it has been named (van Dijk et al. 2000), has created a sense of urgency to locate remaining turtle populations and to evaluate their status before they experience further declines or become extinct. Yet, field research and biological data collection represent the greatest costs associated with biodiversity studies and species conservation (Peterson et al. 1998; Stockwell and Peterson 2002). Therefore, determining turtle species' status in an efficient, economical manner is a critical priority and one that may benefit from the applications presented in this dissertation. This research was a *proof of concept* to examine the efficacy and potential application of GARP ecological niche model, GISc and remote sensing to turtle conservation.

CHAPTER II

THE BOG TURTLE

2.1 Geographic Distribution

The bog turtle, *Glyptemys muhlenbergii*, is North America's smallest and most secretive turtle (Figure 1). It has an average size of 3 to 3.75 inches in straight-line carapace length, with distinctive bright yellow to orange patches on its neck (Herman 2003). Their populations can often be comprised of less than 20 individuals that exist in small, isolated patches of wetland habitat (Buhlmann et al. 1997).

Figure 1. The bog turtle, *Glyptemys muhlenbergii.* (Photo by Dennis W. Herman)

There are two distinct populations of bog turtles separated by an apparent 250-mile disjunct: the northern population, which ranges from New York and

Massachusetts south to Maryland, and the southern population, which ranges from southwestern Virginia and southward to northern Georgia (Ernst et al. 1994) (Figure 2).

Bog turtles have been located in five southeastern states: Virginia, North Carolina, South Carolina, Georgia and Tennessee. Herman (2003) noted that there were less than 1000 individual turtles in the southeast, and that 57% of this total was located in North Carolina alone. There is an ongoing effort to identify the existence of bog turtle habitats and the status of bog turtle populations in each state. Bog turtle presence/absence

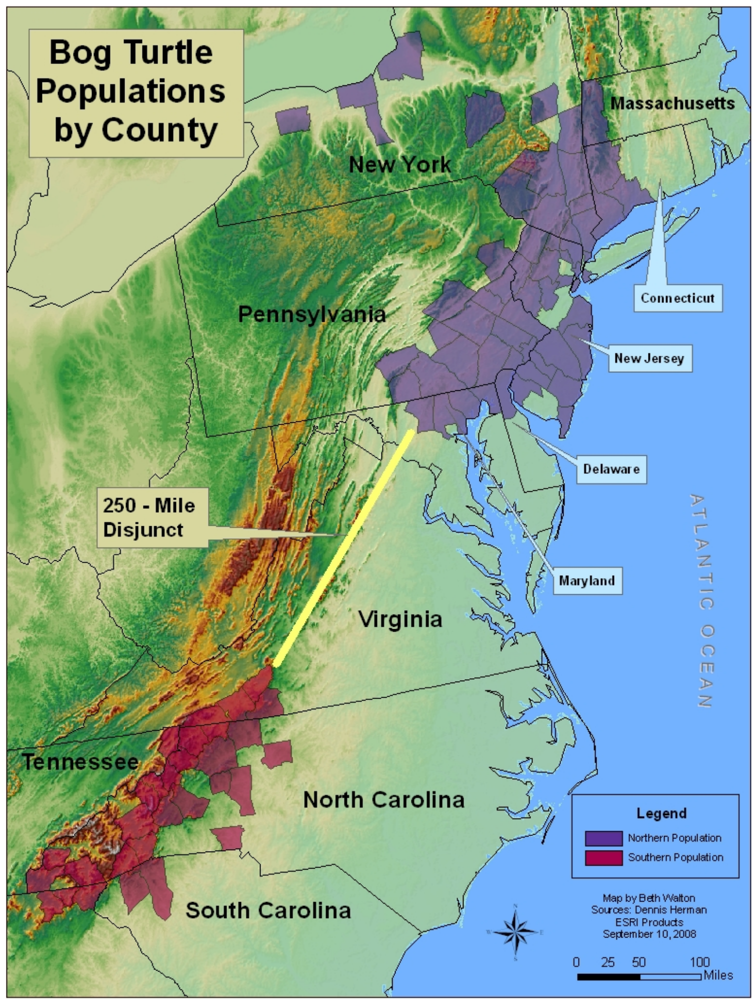

Figure 2. Geographic distribution of the bog turtle. The northern population is separated from the southern population by an apparent 250-mile disjunct.

surveys are conducted in an effort to identify and protect new and potential habitat sites.

2.2 Conservation Status

The northern bog turtle population was listed as “threatened” under the Endangered Species Act (ESA) (1973) in 1997 indicating a loss of habitat as the primary cause of decline (USFWS 1997). The southern population, however, received a lower threat classification and was listed as “threatened due to similarity of appearance” as there was a paucity of baseline population data, and the overall population status and geographic distribution was unknown. Yet, while the southern population did not have sufficient data to warrant the same level of protection as the northern population, it was determined that it would be too difficult for officials to distinguish between them when enforcing the law since the two populations share many phenotypic traits.

Researchers have been conducting ground surveys throughout the southern range since the 1997 ruling in an attempt to locate additional bog turtle populations and to evaluate the overall population status. Conservation management strategies were implemented for some bog turtle populations where landowners were responsive to partnerships with local organizations and agencies, such as Project Bog Turtle and the North Carolina Wildlife Resources Commission.

Locating additional bog turtle populations and habitats has been hampered by a number of obstacles. First, the bog turtle’s small physical size

(3—3.75 inches) and its secretive nature make discovery particularly difficult, even for seasoned field biologists. Bog turtles burrow down in the mud or hide under tussocks in these wetland habitats. Field surveys are conducted by wading through wetland areas using visual searches and probing sticks; a soft tap on the carapace with a probing stick will result in a hollow intonation.

Secondly, many bog turtle habitats are experiencing different degrees of succession. Wetlands left to their own devices will eventually transition to a woodier ecosystem that consumes available water resources. While it is believed that bog turtles will migrate to more open, sunny wetland areas (Herman 2003), many wetlands in successional transition still contain relict populations.

Third, different farming practices produce a range of environmental conditions. Some wetlands have been drained and tiled in years past to put the land into more productive use. These areas will eventually fill in again as clay and silt deposits clog the drainage system resulting in a wetland once again. Other farming practices mow the surrounding fields, yet leave these wet areas undisturbed; still other farming practices merge livestock operations with wetland areas as livestock may enhance the environmental conditions that are favored by bog turtles. These varying environmental conditions make it difficult to discern bog turtle areas by field experts as well as those who utilize GIS, remote sensing and ecological niche modeling technologies.

2.3 Habitat Preferences

Bog turtles are endemic to a narrow range of wetlands in the eastern United

States referred to as *fens*. The defining characteristic of a fen is that its dominant source of wetness is groundwater seepage as opposed to significant inputs from precipitation or runoff from adjacent areas (Bedford and Goodwin 2003). These types of wetlands fall under Cowardin's *palustrine* wetland definition which is described as,

> all nontidal wetlands dominated by trees, shrubs, persistent emergents, emergent mosses or lichens, and all such wetlands that occur in tidal areas where salinity due to ocean-derived salts is below 0.5 ‰. It also includes wetlands lacking such vegetation, but with all of the following four characteristics: (1) area less than 8 ha (20 acres); (2) active wave-formed or bedrock shoreline features lacking; (3) water depth in the deepest part of basin less than 2 m at low water; and (4) salinity due to ocean-derived salts less than 0.5 ‰ (Cowardin et al. 1979, pg. 23).

Additionally, *farmed* wetlands are defined as those,

> where the soil surface has been mechanically or physically altered for production of crops, but where hydrophytes will become re-established if farming is discontinued (Dahl 2000, pg. 74; Cowardin et al. 1979, pg. 11-12).

Fens are also referred to as wet meadows, bogs or meadow bogs and are often found in seepage slopes (Figure 3a) or terraces along the headwaters of small or medium size streams. It is believed that groundwater flowing into these wetland sites assists in the maintenance of more constant temperatures which are cooler in summer and warmer in winter months than surrounding air and surface water temperatures (Amon et al. 2002). Frederick (1974) noted that the waters in Cedar Bog, Ohio never froze and that steam could be observed above the fen during winter months. Additionally, average maximum soil temperatures

Figure 3. Bog turtles prefer a particular type of spring-fed wetlands referred to as *fens:* a) Fens are often visually distinct and sometimes found in seepage slopes; b) Fens are dominated by hydrophilic plant species; c) Iron compounds are reoxidized creating a rusty color; and d) Bog turtles are frequently found in wetlands located in pasture settings. (Photos by Roy Stine)

were 8.88° C cooler in the six warmest months than in nearby soils. These fens may provide a unique environmental setting that allows bog turtles to over-winter in habitats that have warmer temperatures during cold extremes as opposed to other terrestrial sites. They may also provide refuge from extreme heat in summer months and provide some measure of protection against predators (Carter 1997).

Fens are *usually* acidic and experience long-term or continual saturation which produces anaerobic conditions. Anaerobic conditions exist when available dissolved oxygen is depleted by microbiological respiration during the process of decomposing detritus and other organic materials. Once available O_2 is depleted, anaerobic conditions ensue where respiration is conducted by the chemical reduction of iron (Fe), manganese (Mn) and nitrates (N). As iron and manganese are chemically reduced, soils take on a gleyed appearance. Gleyed soils in these fens are typically greenish-blue in color with low chroma values ($\leq$ 2) (United States Department of Agriculture 1975; Munsell 1990).

Wetland perimeters are often visually distinguishable from surrounding areas by their hydrophilic vegetation which is taller and often greener than vegetation of adjacent areas. The fens located in North Carolina are dominated by hydrophilic plant species such as sedges (*Carex* spp.), rushes (*Juncus* spp.), a variety of grasses and vines, and peat mosses (*Sphagnum* sp.) as a common ground cover (Figure 3b). In unmanaged areas where ecological succession is occurring, it is common to find the fern/graminoid plant community transitioning to woodier species such as red maple (*Acer rubrum*), tag alder (*Alnus serrulata*), and tulip poplar (*Liriodendron tulipifera*) (Herman 2003).

In some sites the presence of oxidized rhizospheres may be observed. This results when specialized hydrophytic plant species transport oxygen from parenchyma cells located in leaves and stems down to the root system. As excess oxygen escapes from root tissues, iron compounds are reoxidized creating a rusty color at the root system (Figure 3c). Bog turtles found in sites with high concentrations of iron oxides often exhibit a rusty appearance or carapace pitting (Herman 2003). Other plants may exhibit adventitious roots or spread their roots just below the soil surface in a zone of aeration (Mitsch and Gosselink 2000; Brady and Weil 2002).

Bog turtles are frequently found in wetlands located within pasture settings (Figure 3d) and it is believed that grazing animals, such as cattle and horses, assist in retaining water in these wet areas (Herman 1999). Herman (1999) noted that "96% of southern turtle sites, with turtle densities greater than 20 individuals, were located in currently grazed or recently grazed sites." Grazers churn the mud, creating pockets that trap water that would otherwise flow out of the area, and bog turtles are sometimes found burrowed down in these pockets. Additionally, grazers reduce vegetation that would otherwise remove hydric resources from the site through transpiration. It is believed that these ecological functions were historically performed by American beaver *(Castor canadensis),* which helped to create the wetland areas, and American elk (*Cervus canadensis*) and American bison (*Bison bison*) acted as grazers to deter the growth of woody vegetation (Schafale and Weakley 1990; Lee and Norden 1994) and eventual ecological succession (Herman and Tryon 1997; Tesauro and Ehrenfeld 2007).

There are inferred risks associated with large hoofed animals sharing the same habitat with a comparatively small species such as bog turtles, and especially with respect to eggs, hatchlings and neonates. One management technique used in recent years is the use of temporary livestock excluder fences during months with high turtle activity, usually late spring to late fall (Walton 2002). Working with landowners who had bog turtle populations on their farms, conservation managers installed temporary fencing around the wetland perimeter. The excluder fencing was removed during late fall when turtle activity is low, through to the early spring months to allow grazers into the site to reduce the vegetation and to churn the mud again (Walton 2002). This strategy reduces woody plant species, retains water resources in the site and maintains the open canopy and sunny habitat conditions that bog turtles prefer.

Open, sunny conditions are essential to bog turtles for thermoregulation and it has been the opinion of noted researchers that successional processes that result in canopy closure is the "most critical limiting habitat factor for the bog turtle" (Nemuras and Weaver 1974; Herman 2003). As habitat areas progress through successional processes with greater volumes of woody plant species and closed tree canopies, it is believed that bog turtles historically migrated to other wetland areas that were still open and sunny (Herman 2003).

2.4 Biology and Ecology

Due to a paucity of long-term studies, there is much about the bog turtle that remains unknown. Long-term studies are essential to the effective stewardship of this species as well as to its continued survival in the future. Yet, long-term studies of this species have only begun in earnest within the past 25 years. Bog turtles, like many other turtle species, have the potential to live very long lives, perhaps 50 years or longer. However, it is believed that few hatchlings survive to sexual maturity due to naturally occurring events such as high mortality in eggs, neonates and juveniles. Natural predators include raccoons, foxes and large birds (Bury 1979; Braun and Brooks 1987; and Herman 2003). Herman (2003) noted that “ants, snakes, moles, shrews and rodents are known to prey upon eggs and neonates” (pg. 36).

Bog turtles are opportunistic omnivores that feed on meadow voles (Herman 2003), worms, slugs, snails, beetles, carrion, berries, and seeds (Beans and Niles 2003). Some studies have suggested that the turtle’s environmental role is seed scarification (Braun and Brooks 1987; Calviño-Cancela et.al. 2007). Scarification occurs when seeds pass through the digestive tract and are exposed to water and gases which soften the seed covering allowing for germination. It is possible that the bog turtle serves this same environmental function with the ingestion and dispersal of seeds.

Female bog turtles typically reach sexual maturity when they are approximately 7 years old and will lay an average clutch of 1 to 6 eggs with a mean of 3 eggs (Herman 2003). While annual reproduction is possible, it is unlikely due to fluctuating food sources and infrequent encounters with a male. It is reasonable to assume that females reproduce every second or third season, or 10 to 15 times over their life span. Extrapolated over a lifetime, a female may lay an average of 30 to 45 eggs, but only a few can be expected to survive to sexual maturity (Herman 1994). Many turtle species have natal fidelity, meaning they return to their birthplace to lay eggs or possibly to mate; while it is not certain that bog turtles share this characteristic, it is possible that they do.

Figure 4. Turtles are sexed by external inspection: adult males (right) have a concave plastron with a longer tail and a more posteriorly placed cloaca than females (left). Photo by North Carolina Resources Commission.

Turtles are sexed by external inspection: adult males have a concave plastron with a longer tail and a more posteriorly placed cloaca than females (Figure 4). Male turtles will not show their secondary sexual characteristics, until they reach about 60-mm straight carapace length; however, this feature may be difficult to recognize with the untrained eye (Herman, pers. comm.). Sex determination in young juveniles is almost impossible.

Age determination is measured by counting the annuli or number of rings within a scute. The innermost ring is the natal scute, the scute formed prior to hatching, and is not counted. Each subsequent ring is counted as one year. Figure 5 shows a turtle whose age is estimated to be 6 years old using this method. Bog turtles lose annuli definition as they get older due to repetitive burrowing into the muddy substrate (Herman 2003).

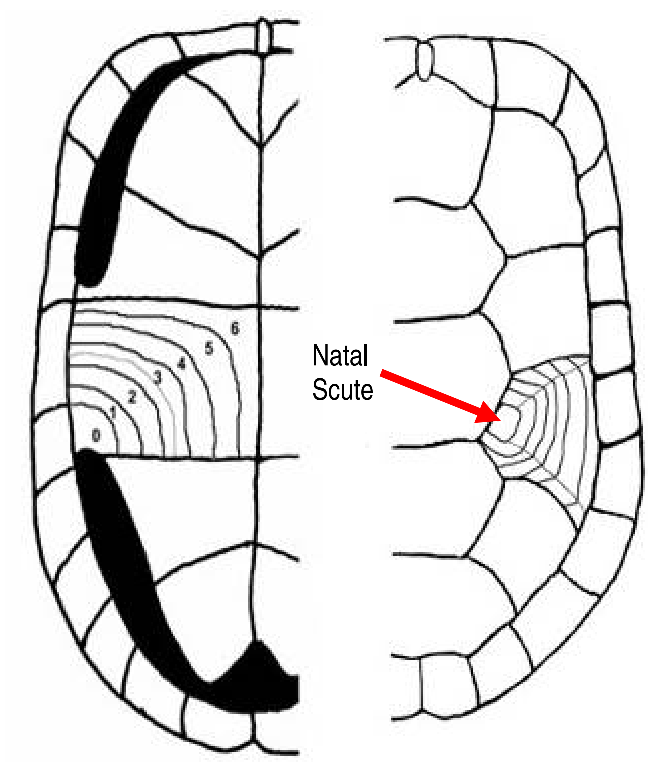

Figure 5. Age determination is measured by counting the annuli. (Modified from Herman)

However, a review and evaluation of this aging technique by Wilson et al. (2003) indicates that there have been no quantitative studies to determine the accuracy and efficacy of this method. Furthermore, it may be more effective to use this technique on younger individuals as annuli definition becomes more difficult to see and distinguish as the annuli become crowded with age. It is also possible that environmental conditions can render the carapace and plastron more difficult to read due to burrowing habits or chemical reactions that take place in the environment. Herman (2003) reported that bog turtle shells can be rendered smooth as a result of constant burrowing into the muddy substrate. He also noted shell "pitting" stemming from iron oxide reactions (Herman 1989; Herman 2003). Thus, shell pitting or smoothing would tend to make it difficult to accurately conduct age assessments.

2.5 Causes of Decline

2.5.1 Habitat Loss

The United States has experienced a significant loss of wetlands since colonial times (Dahl and Allord 2008), but nowhere has this loss been greater than in the Southeast. According to Scott (2005), the Southeast experienced a greater loss of wetlands during the period of 1950 – 1970 than the rest of the United States. Hefner et al. (1994) state, “North Carolina stood out among all southeastern states with an estimated loss of 1.2 million acres in palustrine forested and scrub/shrub wetlands” (pg. 5). Additionally, South Carolina lost 61,000 acres and Tennessee lost 25,000 acres; both states are known to have bog turtle populations.

Small, isolated wetlands play an important role in bog turtle metapopulation dynamics as sites for occasional migration, recolonization, refugia or conservation relocation programs. The loss of these wetlands results in greater migratory distances between wetland sites thus reducing the rate of successful mating opportunities, recruitment and gene flow.

2.5.2 Habitat Loss and Development in Ashe County, North Carolina

Ashe County, North Carolina has experienced a significant increase in development since the 1970s as it has become a popular year-round destination for vacationers, retirees, and the location for recreational homes. Located in the extreme northwest corner of the state of North Carolina, Ashe County lies in the Blue Ridge Mountains and has scenic vistas, two reaches of the New River, state parks and nature preserves. Fall brings thousands of visitors marveling at the colorful foliage and enjoying apple harvests. The relatively mild winters and summers serve as convenient mountain retreats for North Carolinians, as well as visitors from neighboring and distant states. Local attractions include Grandfather Mountain, a nature preserve and wildlife sanctuary that is home to a number of threatened, endangered, and rare species. The Blue Ridge Parkway and the Appalachian Trail cross through Ashe County. The North Fork and South Fork of the New River attract fishing and canoeing enthusiasts while Mount Jefferson State Park, New River State Park, Stone Mountain State Park (all located in North Carolina), and Grayson Highlands State Park in Virginia offer beautiful hiking trails and camping areas. Blowing Rock is a favorite upscale destination that has a variety of bed & breakfasts, novel restaurants, and shops selling unique, local artisan works.

These features have also attracted a significant increase in new building construction and in the real estate market in Ashe County and the surrounding area. As Figure 6 shows, new building construction activity rose from $10 million in 1990 to $80 million in 2007 with similar increases in both the number and size of buildings, as well as multi-family units (United States Census Data). These development pressures have resulted in a significant decrease in the number of wetlands on private property.

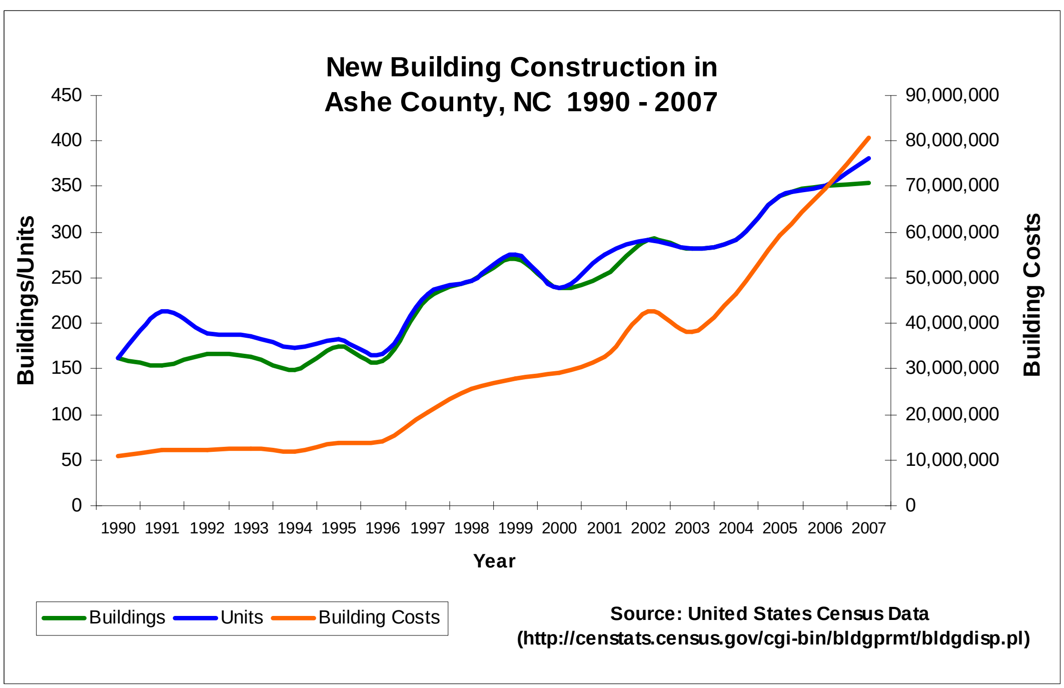

Figure 6. New building construction activity rose from $10 million in 1990 to $80 million in 2007.

Most of the mountainous wetlands in North Carolina are privately held and have been tiled and drained over the years to put the land into more economically productive use. Researchers must use caution when trying to gain access to private property to search for bog turtles as many landowners are leery of government officials and the perceived threat of interference with their property rights. Some landowners deny property access and are resistant to further communication attempts. If bog turtles are found on the property, many landowners are fearful of losing their landowner rights and are reluctant to remove even a few acres from production due to the resulting financial strain.

Through Project Bog Turtle, long-term relationships were established with some landowners and the USFWS provided funds to lease bog turtle wetland sites from them. Modeled after traditional rural land-lease agreements, landowners were paid to leave wetlands unaltered and to allow access to the site for conservation management and research purposes. Several landowners found these leases beneficial as they were protecting an endangered species on their property without sacrificing potential income in the process (Walton 2002). As a result of this program and efforts by researchers to maintain open, friendly relations, many landowners have emerged as important, enthusiastic partners in bog turtle conservation efforts. However, funding from the USFWS for these land-lease agreements has been depleted, and most contracts based on a financial exchange have expired or will expire within the next few years.

In addition to development and agriculture pressures, research conducted by Tesauro and Ehrenfeld (2007) shows a significant decline in small-scale dairy operations since the 1970s. Without the service of livestock to churn the mud and keep these wetland areas open and sunny, ecological succession begins to proceed resulting in woodier vegetation, closed canopies, and thus altered vegetation composition and environmental conditions that don't support bog turtle populations. Many family-owned farms are sold to developers as heirs to the land are no longer available to maintain farming operations, and inheritance taxes prevent them from holding onto the property.

2.5.3 Consumption by Humans and Pet Trade Demands

There is currently a world-wide decline of turtle species and until recently, most scientists would have attributed the global decline to habitat loss, alteration and degradation, and pet trade demands. While these issues remain paramount to turtle species declines, in a relatively few short years these causes have been equaled, and perhaps even surpassed, by the insatiable demand for turtles in China. Taken separately, these causes of decline would prove difficult to mitigate; collectively they may drive many turtle species to extinction (Behler 2000; Rhodin 2000; van Dijk 2000; van Dijk et al. 2000). Used as a preferred food source and in traditional medicines in Asia, the trade in turtles is measured at more than 10—12 million individual turtles per year (Salzberg 1998; Rhodin 2000; van Dijk 2000; van Dijk et al. 2000) creating an urgent call to locate and protect remaining turtle species. Yet, as Gibbons (2000) states,

> . . . the means of determining a species' conservation status is a rigorous and time-intensive process, and therefore counts of "officially" recognized endangered and threatened species are likely to grossly underestimate the actual number of imperiled species (pg. 653).

China's endemic turtle species have been largely depleted and many species may be extinct (Altherr and Freyer, 2000; van Dijk et al., 2000). Other countries around the world are experiencing a depletion of turtle species to satisfy the current demand from China (Behler 1997).

North Carolina turtle populations were dealt a harsh blow when the North Carolina Wildlife Resources Commission's Nongame Advisory Committee reported that commercial collection and harvesting of various turtle species (predominantly aquatic species) in the state had gone from 460 turtles in 2001 to more than 23,000 in 2002. North Carolina is to be commended for having progressive legislators who sprang into action enacting new legislation that became effective July 2003 limiting the number of turtles a collector may possess to fewer than five (North Carolina G.S. 113-333).

While these measures will certainly help to restrict the number of turtles taken after the bill was passed, it will do little to mitigate the damage that has already been done. It has been suggested that the loss of even one reproducing female may have long-term consequences for the local population. The loss of thousands of turtles in a relatively short period of time may be a catastrophic impact to North Carolina's turtle populations and may be evident for many generations to come. It is now believed that the turtles taken from North Carolina were probably headed for European pet trade markets or the food markets of China (Herman pers. comm.; Altherr and Freyer 2000).

The impacts to bog turtle populations stemming from human consumption and pet trade demands are far less significant than the loss of suitable habitats; however, researchers and concerned landowners remain vigilant in protecting the specific whereabouts of habitats that currently support bog turtles. Likewise, this dissertation will be restricted in terms of precise locations that would convey information as to the whereabouts of known or suspected bog turtle habitats. Specific information will be shared only with authorized personnel.

2.6 Field Investigations and Presence/Absence Surveys

The current methodology for bog turtle population discovery is field surveys whereby researchers investigate potential habitat sites with the hope of detecting and protecting remnant populations. In North Carolina and much of the bog turtle's southern range, this involves countless hours of driving around the countryside conducting visual inspections of the landscape. If potential wetlands are observed, the landowners are contacted to request permission to explore on their private property. Once (and if) permission is granted, then many more hours are spent within the site attempting to establish bog turtle presence or absence status. This process, and others like it, is labor intensive and fiscally exhaustive (Vogiatzakis 2003; Rushton et al. 2004), and in cases where there are reduced population numbers, the rates of discovery success are often marginal.

Additionally, road surveys, by nature, are restrictive in scope as they typically only encompass zones that are adjacent to road networks, thus neglecting other areas outside the visual range and across the physical landscape. Therefore, what was needed was a method of investigation that would minimize search time, provide a comprehensive examination of the landscape and optimize the use of labor and financial resources.

2.7 Protecting Endangered Species

The field of conservation biology has experienced a shift in population ecology paradigms. MacArthur and Wilson's (1967) island biogeography model defined the species-area relationship stating that larger islands would support more species than smaller islands and a state of dynamic equilibrium would be attained as new species immigrated and filled ecological niches previously filled by a species that had become extirpated or extinct. This model was widely accepted and evolved to include national parks, wildlife refuges and nature reserves with an inherent concept that conservation measures should be applied to large, contiguous habitats and those currently occupied by species of concern. However, a new paradigm, the metapopulation paradigm, emerged in the 1990s. A metapopulation is a set of local populations contained within a network of habitat patches that allows some migration from one population to another (Levins 1969; Hanski and Simberloff 1997). As opportunities for preserving larger, contiguous habitats became more and more infrequent, and development pressures fragmented and degraded existing habitats, conservation biologists began to emphasize the importance of not only habitats currently occupied by species of concern, but of disjunct habitats and corridors that could serve as sites for occasional migration, recolonization, refugia or conservation relocation programs.

When it comes to protecting imperiled species, conservation biologists, government agencies and other stakeholders are characteristically burdened with a perpetual need for financial resources to sustain research efforts and conservation measures while racing against a ticking clock. Unfortunately, many species simply do not have the genetic ability or the resource of time to adapt to an ever-changing and often degraded environment, and as a result the fate of these species and their continued survival is uncertain (Primack 1995). This is particularly true for specialist species, such as the bog turtle, that act as resource “sinks” in terms of labor and financial investments. A specialist species is one that has a narrow range of environmental tolerances and is usually a species that can extract necessary resources more efficiently than other species in the local environment (Withgott and Brennan 2009). Unlike generalist species that can obtain alternative resources, specialist species are vulnerable to environmental disturbances, especially without the benefit of time to allow for adaptation to new conditions. Thus, when working with endangered species, it is critical to locate potential or additional habitats, monitor existing populations and establish conservation measures, all while working with limited time and financial resources.

This dissertation research provided a unique opportunity to incorporate geographic information science (GISc) with conservation efforts. Multivariate analysis was used to define the geographic and environmental space, and to describe the environmental parameters and physiological tolerance ranges in wetland areas that are known to support bog turtles. Defining the ecological niche of this species and the associated tolerance ranges will be beneficial to researchers seeking additional habitat areas.

It is possible that these cutting-edge, state-of-the-art technologies could become the keystone of conservation strategic planning that propagate new methodologies across many disciplines. Conservation biology, a crisis discipline by nature, may soon be equipped with modern technologies to address issues of species declines with greater efficacy and positive results. In turn, this information will serve to assist researchers, land managers and policy makers in setting conservation priorities for many endangered chelonian species, including the bog turtle.

CHAPTER III

ECOLOGICAL NICHE MODELING

3.1 Historical Foundations of the "Niche" Concept

> *How strange it is that a bird, under the form of a woodpecker, should have been created to prey on insects on the ground; that upland geese, which never or rarely swim, should have been created with webbed feet; that a thrush should have been created to dive and feed on sub-aquatic insects; and that a petrel should have been created with habits and structure fitting it for the life of an auk or grebe! and so on in endless other cases. But on the view of each species constantly trying to increase in number, with natural selection always ready to adapt the slowly varying descendants of each to any unoccupied or ill-occupied place in nature, these facts cease to be strange, or perhaps might even have been anticipated.*
>
> — Charles Darwin (1859), *On the Origin of Species*

Charles Darwin (1859) understood that in a species' "struggle for existence," its survival depended upon its ability to extract and utilize available resources in the local environment more efficiently than its competitors. While the term "niche" hadn't yet been coined, Darwin understood that to survive, each species had to adapt as local conditions changed, and they had to continually compete successfully in order to ensure their survival and the viability of their future generations. And just as organisms evolve to provide ecological services and fill particular positions in the environment, the term "niche" has evolved to provide greater clarity as our understanding of biology and ecology has broadened. Grinnell (1917) is credited with coining the phrase “niche” (pronounced neesh, nitch or nish) and he defined it in terms of a physical ecological address, a place in the environment where all the conditions were conducive for a particular species to thrive. Elton (1927) expanded this definition to include the species’ ecological function or role in the local interconnected community. Hutchinson (1957) described the ecological niche as a composition of an *n*-dimensional hypervolume based upon the numerous environmental variables and pursuant biotic and abiotic conditions and relationships that define an organism’s position within an environment. According to Hutchinson’s definition, an organism’s environmental position and role can change from one geographic area to another, and competition for some niches may be greater depending upon the species involved and the degree of overlapping competition for resources. He defined the *fundamental* niche as all potential *physical* locations regardless of, and in the absence of, competition from other species. The *realized* niche, according to Hutchinson, is a niche that is currently occupied by a species and it includes all interactions of those species that are in direct competition for available resources.

MacArthur’s (1958) work with warblers described how environmental niches may overlap with *niche partitioning* where species evolve to uniquely exploit available resources by methods that are more efficient than other species or by means that other species cannot. This allows for multiple species overlap in the environment, continued species fitness through adaptations, and survivorship with each species extracting necessary resources.

3.2 Using GISc Technologies to Identify Environmental Niches and Habitats

In recent years, conservation biologists have discovered the application of GISc concepts to model species habitats (Braunisch et al. 2008), population distributions (Peterson and Kluza 2005; Ortega-Huerta and Peterson 2008), the spread of disease (Blackburn 2007; Neerinckx et al. 2008) and the spatial distribution and predictions of invasive plant species (Peterson and Vieglais 2001; Peterson et al. 2003). A GIS allows researchers to compile various layers of spatially referenced digital data that are measured representations or proxies of real-world conditions. Spatial analysis allows scientists to map metapopulation habitats, spatial distributions and in some cases to monitor individual movement of larger animal species between habitat patches using radio telemetry, global positioning system (GPS) devices and/or *direct* remote sensing (Davis et al. 1990; Aspinal 1995; Akçakaya 1996; Scott et al. 1996; Chen and Peterson 2000; Peterson et al. 2002; Venkataraman et al. 2002; Turner et al. 2003 Venkataraman et al. 2005, McNaulty et al. 2008).

Some species, however, fall below the threshold that qualifies them for spatial monitoring, which is usually the result of a species' small physical size (Ramanujan 2004), secretive nature (Stockwell and Peterson 2003), or rarity. In such cases it may be more effective and economical to use *indirect* remote sensing techniques to establish a habitat fingerprint or signature and concentrate search efforts in areas that meet the physical and biological requirements of the species' habitat (Turner et al. 2003) or environmental niche.

Remote sensing offers the ability to analyze satellite imagery or aerial photography to determine spectral signatures emitted by objects contained within an area of interest on the Earth's surface. Spectral signatures are a measure of electromagnetic energy reflected or absorbed by objects on the Earth's surface, not only in the visible ranges of blue [45 – 52 nanometers (*μ*)], green (52 – 61 *μ*) and red (63 – 69 *μ*), but also in the near infrared (78 – 90 *μ*), mid-infrared (1.55 – 1.75 *μ*) and mid/short infrared (2.09 – 2.35 *μ*) spectral ranges. Additionally, thermal infrared bands (10.40 – 12.50 *μ*) can detect heat signatures. Using multispectral bands, a signature of the exact biological and environmental assemblages supporting a given species in one area can be used to discover additional habitat areas that have the same fingerprint (Turner et al. 2003) and potential capacity for support.

These spectral signatures, however, are scale dependent: coarse data, such as ≥ 1 km resolution data, are best suited for identifying habitats for generalist species that can survive in a broad range of environmental conditions. Pixels are assigned spectral values based upon the dominant percentage of the landscape features that they represent (Chang 2008). For example, if a 1 km area contains 35% agricultural field, 30% forest stand, 25% wetland area, and 10% water, the representative pixel will be classified as 100% agricultural field as this is the dominant land classification. Thus, small, isolated wetlands can go unrecognized in the pixel classification matrix (Jensen 2005). Hyperspectral imagery, such as the National Aeronautics and Space Administration’s (NASA) Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) or Hyperion satellite imagery, would do a better job of discriminating these wetland areas as the number of spectral bands captured can range between hundreds to thousands, but the costs of this type of imagery acquisition are prohibitive.

Another method of identifying suitable habitats is ecological niche modeling. This methodology employs statistical algorithms to merge spectral signatures with environmental proxies to identify potential habitat areas. Environmental data layers, such as vegetation composition, soil properties, elevation, temperature and precipitation can be used in conjunction with presence data points to assist in identifying specific properties in terms of environmental tolerance ranges that contribute to habitat suitability. This is particularly beneficial when attempting to identify rare habitat areas for specialist species. Areas identified as being suitable habitat and that fall within these environmental ranges can then be investigated for the presence or absence of the species of concern.

Using these new technologies, for example, Raxworthy et al. (2003) created an ecological niche model using a GIS that incorporated satellite imagery and environmental characteristics with records of occurrence (both current and historical) to predict the distribution of rare chameleons in Madagascar. Upon investigating three areas with the highest predictive success rates, *seven new species* of chameleons were identified (NASA website; Raxworthy 2003; Ramanujan 2004).

3.3 Theoretical Foundations of Ecological Niche Models

Soberón and Peterson (2005) proposed a theoretical foundation when using ecological niche models in estimating species distributions. Building on Hutchinson's (1957) definitions of *fundamental* and *realized* niche, these authors theorized that scientists should investigate the intersection of three basic components: abiotic (A) environmental factors (such as temperature, precipitation, pH, organic matter, etc.) that will determine a species' *fundamental* niche, but largely ignores the biotic components; inherent biotic (B) factors (such as predator/prey relationships, natural selection adaptations, and mutualistic/facilitator relationships) based on species occurrence records defining the *realized* niche; and range of movement (M) limitations due to physiology (physical characteristics which allow or restrict the ability to disperse) or environmental barriers (such as mountains, dams, water bodies, etc.) that restrict dispersal. It is the intersection of these three variables in which scientists will theoretically find the perfect (P) conditions and greater likelihood of prediction success in the real world (Figure 7).

Araújo and Guisan (2006) argue that Hutchinson's (1957) definitions of fundamental and realized niches are based upon competitive exclusion in smaller spatial extents. Yet in reality, competitive species can and do co-exist through resource partitioning on a local level, as described by MacArthur (1958), and they can certainly exist within larger geographic extents without direct competition for resources. Araújo and Guisan propose that when using ecological niche models, researchers should adopt a new niche definition as proposed by Chase & Leibold (2003, p 19):

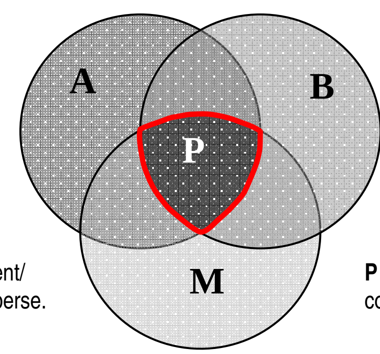

Figure 7. Theoretical foundations of ecological niche models. It is the intersection of abiotic (A), biotic (B) and movement (M) in which the perfect (P) environmental conditions exist. (Adapted from Soberón and Peterson 2005)

> *the environmental conditions that allow a species to satisfy its minimum requirements so that birth rate of a local population is equal to or greater than its death rate.*

Araújo and Guisan (2006) suggest that niche modeling based strictly upon environmental factors produces *potential habitats;* yet when niche and spatially explicit layers are combined, the niche model yields the "*potential geographical species distribution*" and this is where GISc technologies and ecological niche modeling merge.

A GIS allows the simulated replication of environmental variables and analysis of the environmental tolerance ranges within which a species is currently able to exist. These environmental variables can be stacked so that each layer is in the same planimetric (*-x, -y*) position with presence data points overlaying the stack. These data layers are then utilized in an ecological niche model which will analyze all the variable values associated with each presence data point. The ecological niche model will ultimately produce a map predicting areas that have the greatest likelihood of replicating the range of conditions as defined by the range of values for all the presence data points. This prediction map is then projected onto geographic space indicating selected areas that very closely resemble the environmental assemblages that support the species under study and thus are areas of potential distribution where additional individuals might be discovered.

It is important to recognize that ecological niche modeling is not an exercise of predicting where a particular species might be found. Instead, it is a process of identifying the environmental conditions that exist where species are known to occur, and understanding that it is these specific environmental and biological conditions that provide the capacity for support. This is often

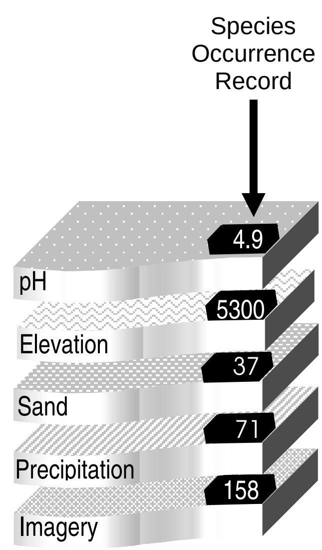

Figure 8. The species' *environmental envelope* will identify the value for each variable at points of presence data.

referred to as the species' "*environmental envelope*" and it will identify the environmental value for each variable at points of presence data (Figure 8). These values can then be analyzed to establish environmental tolerance ranges for each variable; descriptive statistics can provide the mean, mode, and standard deviation for all variables used in the model. These variables can also be analyzed to determine which ones have more influence on habitat suitability and whether there are specific limiting conditions beyond which the species cannot exit (Peterson and Kluza 2005; Pearson 2007). It is also important to emphasize that environmental and biological variables are estimations of real-world conditions that often have inclusions (e.g. soil and vegetation types), interpolations (e.g. estimates of temperature and precipitation values between weather recording stations), omissions (a lack of information) and generalizations (e.g. loss of real-world detailed information).

3.4 Environmental Space vs. Geographic Space

According to Pearson (2007), the assemblage of environmental proxies in a GIS is a simulation of the biological and environmental conditions found in the real world. The incorporation of these proxies into an ecological niche model produces predictions of *potential* habitats in a theoretical *environmental* space or area of interest that correlates with Hutchinson's *n*-dimensional niche (1957), where *n* is the number of environmental variables (layers) used. After completing the modeling process, these predictions can then be projected into *geographic* space in the real world and investigated for accuracy.

Pearson (2007) explains that ecological niche models are designed to predict two different outcomes. Some models can be used to verify the *actual* distribution of a species, which is every area that the species currently exists, whether these areas have been discovered or not. Alternatively, the model can produce *potential* habitats which are suitable habitat areas that may be used for migratory corridors, refugia, head-starting or relocation programs, regardless of whether or not they are currently occupied by the species of interest. These areas may represent habitats that were historically occupied, but from which the species has been extirpated, areas that the species could migrate to at some point in the future, or suitable areas for repatriation or head-starting programs (Figure 9).

3.5 Genetic Algorithm for Rule-set Production (GARP)

The Genetic Algorithm for Rule-set Production (GARP) (Stockwell and Peters, 1999) is an ecological niche model that utilizes spatially referenced biological and environmental data layers to predict the potential geographic distribution of a given species. Essentially, the program looks at all the environmental and biological variables as genetic combinations. Just as is true in biological situations where many deleterious genetic combinations do not survive, the GARP modeling program will delete the combinations that include invalid variable (genetic) combinations. Based upon known habitats with successful combinations of these variables, GARP predicts which sites are more likely to support a population by analyzing the probability of class membership (presence or absence) for variable combinations (Vogiatzakis 2003).

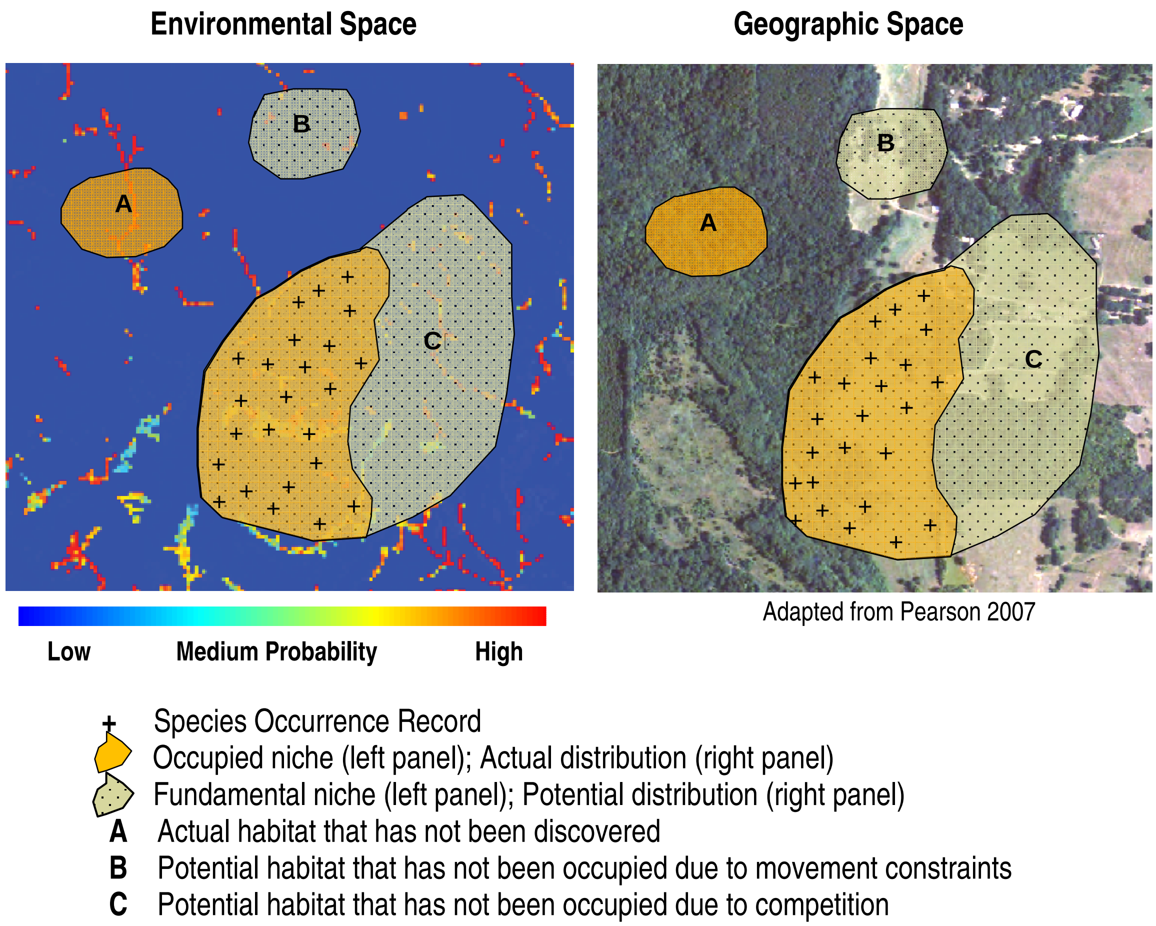

Figure 9. Environmental vs. geographic space.

The GARP program is initiated with a set of “training” data consisting of one-half of the total number of presence data points from which the program “learns” successful combinations of environmental and biological characteristics. A second set of data points, representing the other half of the presence data points and known as “test” data, are used to determine the accuracy of the model's predictions. A successful model is one in which the program can accurately predict where the test data points are located, as well as other areas that are environmentally similar and which represent areas of potential habitats in geographic space (Peterson 2001, Stockwell's GARP User's Manual, Anderson et al. 2003; Pearson 2007).

The GARP program will consider all the environmental and biological characteristics, as well as proximity to neighboring cells to predict potential habitat patches. The program is stochastic in nature which correlates with genetic diversity and biodiversity, thus each surviving model will have successfully persisted through a large number of iterations (~ 1000). Statistical rules will continually improve model predictions until a single model with the highest predictive success rates is developed. No two models, however, will be exact due to this element of randomness (Stockwell's GARP User's Manual, Anderson et al. 2003; Pearson 2007).

Once the iterations are completed, a raster image is generated using a color ramp to indicate areas of higher probability of potential habitat sites. The top models with the highest levels of predictive accuracy are used in a GIS to determine areas of overlap, thus indicating areas of highest probability of habitat suitability. The final model will delineate areas that are environmentally similar to areas that currently support bog turtles and thus may be explored to determine suitability for bog turtle habitat and the presence/absence of bog turtles. It may also be possible to measure distances between habitat patches and infer migratory routes and habitat corridors.

Ecological niche models have inherent predictive limitations. Common errors arise from:

1. presence only data biases where few or no absence records exist;
2. inconsistent data collection methods;
3. false positives (commission) indicating a species presence when it is not; and
4. false negatives (omission) indicating a species absence when it is present (Loiselle et al. 2003).

GARP incorporates a number of statistical and operational methods (jackknifing, bootstrapping, and heuristics) to largely circumvent these issues (Stockwell and Peters 1999). Jackknifing is a numerical resampling technique producing new data sets by eliminating one observation from the original data set at a time, and thus the number of new data sets (each with n -1 total observations) is equal to the number of observations in the original data set. This technique allows for calculating variance within a data set. Bootstrapping is a statistical method using repeated random sampling with replacement from an original sample to provide a collection of new pseudoreplicate samples from which sampling variance can be estimated. Heuristics is defined as a rule of thumb, generally based on expert experience or common sense rather than an underlying theory or mathematical model that can be incorporated in a knowledge base and used to guide a problem-solving process (Peterson and Cohoon 1999).

3.6 Previous Work

Previous research conducted by this author entailed the use of remote sensing to model habitats and population distributions of the bog turtle (Walton 2006). A subset of a Landsat 7 Enhanced Thematic Mapper+ image and color-infrared (CIR) digital aerial photographs were used to analyze a portion of Ashe County, North Carolina, where bog turtle habitats and wetlands occur. The results showed that publicly available data may not be suitable for detecting small, isolated wetlands across the landscape due to heterogeneous landscape features, low spatial resolution and temporal discrepancies of the images, and the inherently poor quality of some of the images. However, it *was* possible to define spectral signatures for wetlands where quality, high spatial and temporal resolution CIR data were available (Walton 2006).

The objective of this dissertation was to determine whether GISc technologies could be used to model small discontinuous wetland habitats similar to those habitats currently occupied by bog turtles using the GARP ecological niche model and high resolution data. High resolution CIR aerial imagery from the National Agricultural Imagery Program (NAIP) served as the basis for the model. Unlike imagery used in the original analysis, these NAIP images were of greater resolution (1 meter) and quality. The study area was within Ashe County, North Carolina which was accessible with a short (~1.5 hour) drive by automobile, and which contains several well-established bog turtle populations.

This research is unique in several aspects: it will represent the first application of ecological niche modeling to a chelonian species; it will be the first known application that uses high resolution aerial photography (1-meter); and it will be the first application in which extensive ground referencing is conducted to measure model accuracy.

3.7 Hypotheses

The GARP ecological niche model has the potential to serve as a flexible, robust conservation tool that will exceed theoretical implications with empirical field results. While theoretical inferences play an important role in the scientific process, empirical data are crucial in developing conservation strategies and policies to afford adequate protection to the species of concern. In the case of the bog turtle, it is imperative to identify the location of potential or occupied habitats as many wetlands are located in areas of rapid development. It is also important to determine bog turtle presence/ absence status and to implement habitat preservation strategies. To this end, this research proposal had three principal goals:

Goal #1: To evaluate the role of GISc, GIS, remote sensing and GARP ecological niche modeling in detecting small, discontiguous wetland areas that may support bog turtles and thus bog turtle conservation efforts. This includes developing an ecological niche model and verifying prediction accuracy with ground referencing. The hypothesis is that the GARP predictive model will detect potential wetland habitat areas and field investigations will support model predictions with newly discovered areas that are potential habitat sites or that may serve as migratory corridors, refugia, or head-starting/relocation program locations. Measurable outcomes will be determined by the presence of hydric soils, wetland vegetation, bog iron, or some combination of these three variables.

While the author holds an endangered species permit, it was not the intent of this research to actively search for bog turtles in any potential wetland areas. Landowner relationships, which are essential to bog turtle management plans, require a lot of time and effort to establish adequate means of communication and trust before a landowner will allow their property to be investigated. Additionally, it will take seasoned investigators and possibly extensive investigations that could endure for more than a year to determine species presence or absence status. Results of this research will be submitted to USFWS authorities and the North Carolina Wildlife Resources Commission which serves as the state agency that administers and enforces rules of the Endangered Species Act in the state of North Carolina. Newly discovered areas will be investigated further to determine whether or not they contain bog turtles. The author's endangered species permit is only for incidental finds and no known habitat sites will be investigated as part of the ground referencing process.

Goal #2: To establish the criteria necessary to determine why bog turtles exist in specific wetlands. This research will identify the environmental tolerance ranges for variables used in this analysis that currently support bog turtle populations within the study area of Ashe County, North Carolina. This information will assist researchers in understanding the environmental extremes in which this species can exist, it will assist in locating potential habitat sites with similar environmental ranges, and it will assist in management strategies. Through future studies, it will also assist in finding additional habitat areas throughout the bog turtle's southern habitat range. However, it should be noted that this study only includes a portion of one county within the bog turtle's southern range. It is possible that there are additional environmental gradients and/or that the environmental tolerance ranges may shift as additional counties are investigated. As such, findings produced by this research may not provide an all-encompassing habitat definition; future studies will be needed to determine if these results and habitat descriptions prevail throughout the southern habitat range.

It is hypothesized that since bog turtles have an intimate relationship with the soils in which they reside, soil and hydrology data layers will play an important role in where bog turtle populations occur. Spectral reflectance values from remotely sensed imagery will provide quantitative data to support model predictions. While researchers are aware of some specific habitat parameters, the GARP model may produce statistical results that will identify environmental and biological range specifics that have not been previously documented. Measurable outcomes of this goal will include descriptive statistics of presence data points to yield a suite of environmental tolerance ranges that currently support bog turtles.

Goal #3: To establish methodologies that will eliminate unsuitable wetland areas where bog turtles do not or will not occur because these areas do not meet the minimal criteria necessary to qualify as a potential wetland habitat. This will include environmental ranges that fall outside the environmental tolerance range as defined by descriptive statistics derived from the presence data.

It is hypothesized that this research will reveal “hidden variables” that would identify where bog turtles are likely *not* to exist due to a lack of suitable environmental conditions to support their biology and ecology. These hidden variables may help to explain habitat requirements that are not easily observable or previously known. It is also possible to identify variable data layers that add less to overall model predictive accuracy and which can be eliminated in subsequent model runs. Measureable outcomes of this goal will be to run the GARP model with a streamlined suite of environmental variable layers and to be able to identify those variables that do not improve model predictive accuracy.

The final prediction maps will be generated with the GARP model and they will be submitted to the USFWS and the NCWRC so that trained field researchers can investigate areas of greater predictive accuracy to determine bog turtle presence/absence status. Though not statistically significant, this hypothesis will be biologically significant if even *one* additional bog turtle (current or historic evidence) is located as a result of model results and subsequent field investigations. Theoretically, even if bog turtles are not found in areas of higher prediction, it is possible that they may have existed in these areas historically and the identification of suitable habitat will add to the scientific knowledge base as areas that may be used for migratory corridors, refugia, head-starting, repatriation or relocation programs.

CHAPTER IV

RESEARCH METHODS

4.1 Data Layers

The purpose of this research was to evaluate the role of GIS, remote sensing and ecological niche modeling in detecting small, discontiguous wetland areas that may support bog turtles and thus bog turtle conservation efforts. This included gathering and manipulating environmental data layers and remotely sensed imagery, developing an ecological niche model and verifying model prediction accuracy with ground referencing.

One of the first considerations for this research was to determine which environmental data layers would be necessary to build the GARP model and which data layers would be most appropriate in the identification of bog turtle habitats. A literature review was conducted to determine what types of data were commonly included in other studies that utilized the GARP ecological niche model, and how many data layers were incorporated into a typical model. While Peterson and Cohoon (1999) caution that the incorporation of too many data layers may cause “overfitting,” or an overestimation of potential habitat sites, many successful applications of the GARP model used a large number of environmental variables. Stockwell and Peters (1999) indicated that most GARP modeling systems used 30 data layers; Raxworthy (2003) used 29 data layers in his model for chameleons; Pearson et al. (2006) used 20 environmental data layers for his work with “cryptic” habitat modeling for geckos; and Blackburn et al. (2005) started with 34 environmental data layers to model the spatial ecology of anthrax (*Bacillus anthracis)*. It was anticipated that with the inclusion of more variables, certain hidden variable associations or emergent properties would be discovered to assist in the detection of obscure mountain wetlands, and to isolate those factors that explain bog turtle selection of certain wetland areas over others. Isolating complex combinations of environmental variables would aid in the location of additional habitat areas with the same suite of variables.

Since bog turtles are found in wetland habitats, it seemed obvious that soil type and soil related properties would be significant in determining wetland areas and bog turtle habitats. Hydric soils associated with hydrophilic plant species located near a water source are likely key components in teasing out small, isolated wetland areas with high resolution data. It was further proposed that habitat orientation and length of daily solar exposure were key factors in determining bog turtle preference for certain wetland areas. Turtles are ectothermic and must use thermoregulation to maintain their body temperatures. An aspect data layer could be a key factor in the detection of potential habitats with sufficient solar radiation. A number of other studies have included precipitation and temperature as key variables in detecting suitable habitat areas (Peterson and Cohoon 1999; Raxworthy 2003; Blackburn et al. 2005; Peterson and Nakazawa 2008) and it was expected that this study would not be an exception. North Carolina has experienced long-term (> 10 years) drought stressors (Weaver 2005; North Carolina State Climate Office; North Carolina Drought Management Advisory Council). Under these conditions the use of color infrared imagery could help distinguish areas of healthy hydrophilic plant species with longer access to hydric conditions.

A data layer that represents predator/prey relationships would be ideal in species modeling. However, bog turtles are exceptionally rare and would only provide an occasional, chance diet supplement to natural predators, such as raccoons, foxes or birds (Bury 1979; Braun and Brooks 1987; and Herman 2003). Alternatively, bog turtles are opportunistic omnivores that can feed in the water or on land and typically forage and feed on a number of prey including meadow vole “pinkies” (Herman 2003; Herman pers. comm.), worms, slugs, snails, beetles, carrion, berries, and seeds (Beans and Niles 2003). Since these prey species were commonly found in abundance in a variety of natural environments, it wasn’t practical to include a data layer that would represent their presence. A complete list of the data layers used in this research can be found in Appendix A. The following is a description of the assembled data layers used in this research.

4.1.1 Aerial Imagery

Previous research has indicated that high quality, high resolution imagery is necessary to obtain distinct spectral signatures of wetland areas across the landscape. That study found that small, discontiguous wetlands in various stages of succession could not be distinguished using low spatial resolution imagery in the form of a Landsat ETM+ 30-meter image (Walton 2006).

For this study, United States Department of Agriculture (USDA) National Agriculture Imagery Program (NAIP) color-infrared (CIR) digital orthophoto quad (DOQ) MrSid mosaic and a natural color DOQ MrSid mosaic for Ashe County at 1-meter resolution were selected. The NAIP acquires color DOQs during the spring and summer crop growing seasons across the continental United States (USDA NAIP) every other year. Natural color imagery was available for 2004, 2006 and 2008; the 2006 imagery was used in this research.

Color-infrared (CIR) imagery was available for the year of 2005 and this imagery was used in this research as well. The CIR imagery was captured in three spectral “bands” from the electromagnetic spectrum. The spectral range is approximately 500-600 nanometers (nm) for the green band; 600-700 nm for the red band; and 700-900 nm for the near-infrared (NIR) band. CIR near-infrared (NIR) wavelengths lack ultra-violet and blue wavelengths which results in a CIR image. Green, healthy vegetation (such as grass) has a high reflectance of NIR wavelengths and appears as bright red; vegetation with very low NIR reflection appear green (such as conifers); non-vegetative green objects with very low NIR reflection appear blue; and blue objects with very low near-infrared reflection appear black (such as water). The spectral reflectance of soil is strongly correlated with moisture content: high moisture content results in lower reflectance (Jensen 2005).

The imagery bands are stored in raster grid data formats with cell values based upon the spectral reflectance (brightness) values of each pixel. The spatial resolution of 1-meter should enhance the ability to identify smaller wetland areas. Principal components analysis (PCA) or single band composites were used to isolate and differentiate wetland areas. PCA is a measure of variability between bands while reducing spectral “noise” incurred from atmospheric distortions or mechanical inconsistencies (Jensen 2005; Shlens 2009) and this variation was presented in the form of a new data layer that was used in the ecological niche model.

4.1.2 Normalized Difference Vegetation Index and Soil Adjusted Vegetation Index

The 1-meter NAIP 2005 CIR digital orthophoto quadrangle (DOQ) MrSid mosaic was used to create a normalized difference vegetation index (NDVI) layer using the following formula per Jensen (2005):

$$\boldsymbol{NDVI = \frac{(NIR - R)}{(NIR + R)}}$$

Where:

NDVI = Normalized Difference Vegetation Index

NIR = Near Infrared Band

R = Red Band

Healthy, green vegetation has a high reflectance value in the CIR band, thus the NDVI would readily distinguish these areas. Due to the relationship with hydric soils, hydrophilic vegetation should appear healthier and greener than surrounding vegetation, especially during times of drought. A SPOT (Satellite Pour l'Observation de la Terre) 1995 10-meter satellite image obtained from the University of North Carolina, Geography Department, data library was used to create an NDVI data layer as well.

North Carolina has experienced overall drought conditions since the 1980s. The North Carolina Drought Management Advisory Council was organized 1992 and given additional drought monitoring and advisory authority in 2003. These drought conditions were particularly severe during 1998—2002 (Weaver 2005) and have persisted ever since with only occasional respite. In response to these conditions, a soil adjusted vegetation index (SAVI) data layer was created from the NAIP 2005 CIR image to minimize soil background conditions, such as soil brightness, while optimizing the spectral signatures of the vegetation. Hydric soils, as stated above, however, should have a low spectral reflectance value. The SAVI layer was used in the event that drought conditions were affecting the upper soil layers, thus introducing some measure of soil brightness. The SAVI layer was created using the following equation as per Jensen (2005):

$$SAVI = \frac{(NIR - R)\quad(1 + L)}{(NIR + R + L)}$$

Where

SAVI = Soil Adjusted Vegetation Index

L = Soil-brightness dependent correction factor (.50) (per Jensen 2005)

NIR = Near Infrared Band

R = Red Band

According to Jensen (2005, p 318),

> *L* is a canopy background adjustment factor that accounts for differential red and near-infrared extinction through the canopy. An *L* value of 0.50 in reflectance space was found to minimize soil brightness variations and eliminate the need for additional calibration for different soils, (Huete and Liu, 1994 as referenced in Jensen 2005).

4.1.3 Elevation Data

A 10-meter digital elevation model (DEM) was obtained from the United States Geological Survey National Elevation Dataset and used to create a slope layer. A percentage of slope from “0 - 9” represented a relatively flat surface and “91 – 100” represented a steep slope. Most wetlands are located in 0 – 10% slope which distinguishes depressional areas receiving hydrologic input from adjacent areas. Wetland *fens,* however, are predominantly supplied with groundwater recharge (Bedford and Goodwin 2003) and this research documents the percentage of slope in known bog turtle wetland areas. The percentage of slope will be used to determine the applicability of slope to bog turtle habitats within the research area.

The DEM was also used to create an “area solar radiation” data layer in ArcMap. The area solar radiation data layer is a representation of the duration of solar exposure during specified times of the year using topographic features and latitudinal position of a local area.

4.1.4 Aspect Data

An aspect data layer provided valuable information regarding the orientation of prime wetland areas for bog turtle habitation. It is proposed that most wetlands utilized by bog turtles are oriented for maximum solar radiation during the daytime where

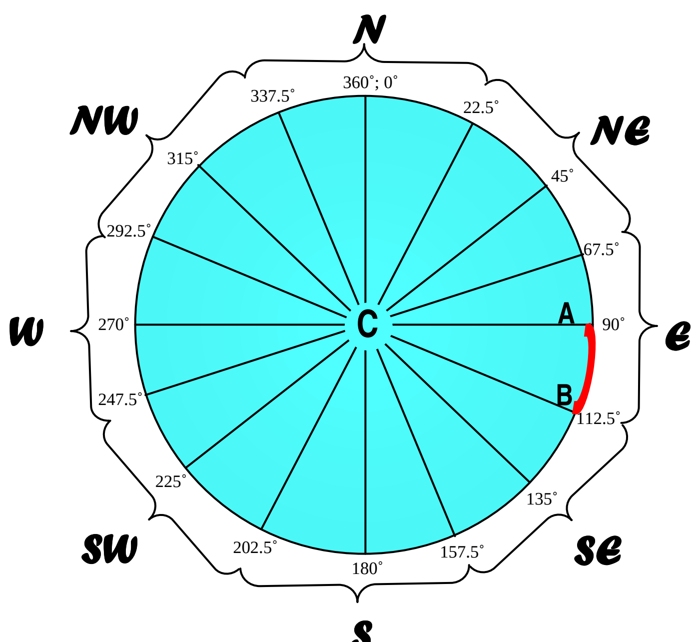

Figure 10. Degrees can be converted into cardinal directions (north, south, east, and west) with subdivisions (northeast, southeast, southwest, or northwest).

bog turtles bask in sunlight to maintain body temperatures. The 10-meter DEM was used to create an aspect data layer. Most aspect data layers are in the form of degrees which can be misinterpreted by modeling systems to be hierarchical in nature so that 355˚ is significantly larger than 0˚ when in actuality they are much closer in value. Degrees can be converted into cardinal directions (north, south, east, and west) with subdivisions (northeast, southeast, southwest, or northwest); however these are classified as categorical data (Figure 10). Degrees can also be converted into radians, which are a measurement of the angles from the center of the circle to points A and B on the circle's radius using the following formula:

$$1^\circ = 1 * \frac{\pi}{180^\circ} \approx 0.0175 \text{ radians.}$$

The DEM was converted from degrees to radians in ArcMap by calculating the cosine and sine using a raster calculator and the following formulae as per Huber (2008): as follows:

Layer 1: **Cos([Aspect of dem] * 3.14159 / 180)**

Layer 2: **Sin([Aspect of dem] * 3.14159 / 180)**

and then converted the degrees to angles:

ATan2(FocalMean([Sin], CIRCLE, 50), FocalMean([Cos], CIRCLE, 50)) * (180/3.141593).

4.1.5 Soils Data

It was important to profile habitat characteristics that are currently associated with bog turtles in order to develop an ecological niche model to predict additional sites. While there is still much to learn, researchers have identified many characteristics that most occupied habitats share in common (Carter et al. 2000; Chase et al. 1989; Ernst et al. 1994; Herman 2003). For example, wetland fens are one of the most dominant habitat types for bog turtles. Thus, it was important to analyze edaphic characteristics that could play an instrumental role in bog turtle population distributions. Bog turtles have an intimate relationship with the soils in which they reside as a result of their physiology, biology and ecology. As such, wetland soil characteristics and associated soil attributes constituted a large number of the layers used in the GARP model.

Soil related data layers were created using the United States Department of Agriculture (USDA) Soil Survey Geographic (SSURGO) data for Ashe County, North Carolina. Individual layers were created using the Soil Data Viewer (SDV) software developed by the Natural Resource Conservation Service (NRCS). Typical parameters for data layer creation included an aggregation method of "dominant component" and a depth range of 0 – 36 inches. Appendix B provides attribute definitions and descriptions of each data layer used in this research as provided by the SDV module. Soil data layers included those shown in Table 1 below:

Table 1. Soil Data Layers Created with the Soil Data Viewer		
Available Water Capacity	Electrical Conductivity	Plastic Index
Available Water Supply (0 – 25 and 0 – 100 cm)	Farmland Classification	Ponding Frequency Class
Bulk Density	Flooding Frequency Class	Sand
Bulk Density 1/3 Bar	Frost Action	Saturated Hydraulic Conductivity
Bulk Density 15 Bar	Frost Free Days	Silt
Calcium Carbonate	Hydric Rating	Slope
Cation Exchange Capacity	Hydrologic Group	Sodium Absorption Ratio
Cation Exchange Capacity (CEC-7)	K-factor Rock Free	Soil Texture
Clay	K-factor Whole	Surface Texture
Depth to Soil Restrictions	Liquid Limit	T-Factor
Depth to Water Table December and July	Organic Matter Content	Water Content 15th Bar
Drainage Class	Permeability Class	Water Content 3rd Bar
Effective Cation Exchange Capacity	pH Rating	

4.1.6 Weather and Climate Data

Weather data were obtained from the North Carolina State Climate Office (www.ncclimate.ncsu.edu) using cooperative observer program (COOP) data for the Jefferson, NC Station (ID 314496), Laurel Springs, NC Station (ID 318694), Sparta, NC Station (ID 318158) Abingdon, VA Station (ID 440021) Troutdale, VA Station (ID 448547) Boone, NC (ID 310982) Wilkesboro, NC Station (ID 319555) and the Ashe County Airport Weather Station (KGEV). Data included minimum, maximum and mean temperature, precipitation, wind speed and wind direction where available.

Additional precipitation and temperature data layers were obtained from the NRCS for the state of North Carolina. The precipitation and temperature data sets used the Parameter-elevation Regressions on Independent Slopes Model, commonly referred to as *PRISM* data. PRISM data sets are created using a DEM and point climate data to generate monthly and annual estimates of temperature and precipitation by means of a “hybrid statistical-geographic approach to mapping climate” (NRCS PRISM Data Sheet, pg. 1; downloaded February 2009). Data layers in this set included annual precipitation, monthly precipitation estimates, and temperature minimum, maximum and average. In addition, one hundred and three year (1895-1997) PRISM high-resolution temperature climate data was obtained from the PRISM Climate Group website (www.prism.oregonstate.edu) and clipped to the spatial extent of Ashe County, North Carolina.

4.1.7 Ashe County Watersheds and Subbasins

Ashe County is located in the upper most north-west corner of North Carolina and is bordered by Tennessee on the west and Virginia on the north. Ashe County is in the Upper New River Watershed and is associated with nineteen subbasins (Figure 11). North Fork and South Fork are the two main branches of the New River which cross the county. The highest point of elevation is 5,208 feet located at the Peak on the western side of the county (USGS topographic maps).

Bog turtle habitats are highly correlated with streams; therefore a number of hydrologic data layers were obtained to denote watersheds, subbasins and stream locations. These data files were obtained from the United States

Department of Agriculture Natural Resource Conservation Service's National Cartography and Geospatial Center and included twelve-digit Watershed Boundary Data at a scale of 1:24,000 and eight-digit hydrologic unit codes (HUCs) at a scale of 1:250,000.

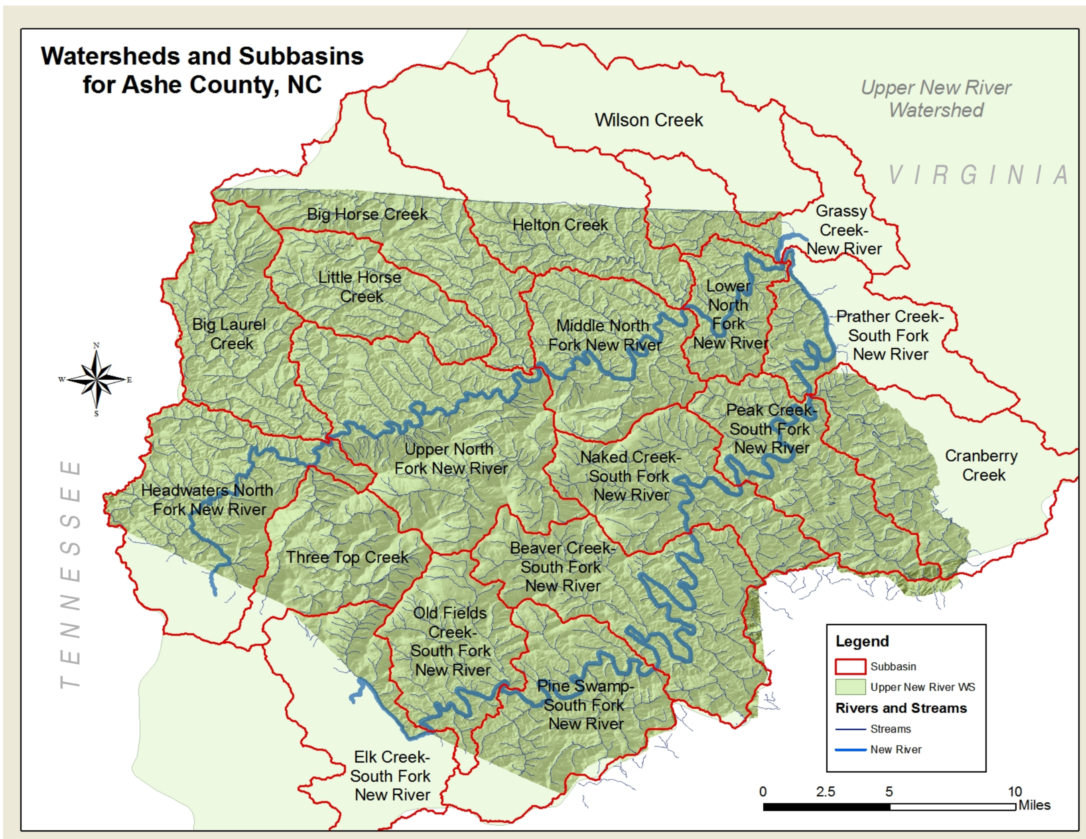

Figure 11. Ashe County watersheds and subbasins

HUCs are a hierarchical mapping system of water containment units developed by the United States Geological Survey. Each HUC has a unique identifier and can range from two to twelve digits in six levels. The first level and first two digits represent the Region; each additional two digits represent the next level and are classified as Subregion, Accounting Unit, Cataloguing Unit, Watershed and Subwatershed respectively. These data layers were used to create hydrologic flow directions, streams and stream orders for analysis in this research.

4.1.8 Wetland Polygons

As indicated previously, this author did not enter any known wetland sites during the course of the research conducted for this dissertation. During the research conducted for the Master of Arts degree, however, the author worked in conjunction with Project Bog Turtle (PBT), the North Carolina Wildlife Resources Commission (NCWRC), The Nature Conservancy (TNC) and the University of North Carolina at Greensboro to obtain data to construct a wetland polygon data layer. An Endangered Species permit was obtained from the NCWRC, the issuing authority for the USFWS Endangered Species Office. A research permit was obtained from TNC as one of the research sites was located on their property.

Dennis Herman, one of the original founders of PBT, provided the coordinates of 13 wetland sites. Six sites were known to support bog turtles and 7 appeared to be suitable habitat, but bog turtles have never been found in these locations. Herman provided a guided tour of the sites on 11-Aug-05 to allow visual inspection, to obtain ground reference information and to take photographs of dominant floral species.

Landowner information was provided by PBT and verified with the Ashe County GIS tax parcel website (http://ashegis.ashecountygov.com/webgis/). Letters were sent to landowners explaining the purpose and nature of the research and requested permission to conduct research on their properties. Information was also requested from the landowners regarding hunting practices on their property; parking preferences; whether they would like additional information on bog turtles; if they wanted to be contacted before a site visit; the preferred method and point of entry into the wetland; and since many of the sites are located in pasture settings, whether or not they had “charging bulls” on the property. Permission was granted from all six landowners with wetlands that currently support a bog turtle population on their property. Only one landowner without a known bog turtle population responded and granted permission.

Another visit was made on 08-Oct-2005 to the properties that granted permission to delineate the wetland areas using a Trimble GeoExplorer CE global positioning system (GPS) to use as ground reference information. The Trimble unit was set to collect data using US State Plane 1983, North Carolina 3200, North American Datum 1983 (Conus), GEOID99 coordinate system, collecting spatial reference points every five seconds. A minimum of four satellites were triangulated at all times during the collection process with a maximum of seven satellites as shown in Table 2. The perimeter of each wetland was walked using the distinct hydrophilic vegetation as a guideline.

Table 2. Data and Readings from Trimble GPS Unit					
Site Name	**Satellites**	**Min # Satellites**	**Max # Satellites**	**Precision**	**Time**
ASHE05	3,13,16,19, 23, 25, 27	5	7	68%	12:41:32p
ASHE09	No Data				
ASHE10	No Data				
ASHE15	7,8,11,19,27,28	4	6	68%	04:00:12p
ASHE18	3,8,13,16,19,23,27	4	7	68%	02:14:00p
ASHE19	1,3,13,16,20,23,25	4	7	68%	11:25:01a
WetlandD	1,14,16,20,25,30	5	6	68%	09:23:10a

A differential correction was performed on the wetland data files using the base station from Conover, North Carolina. Corrections were made to the numbers and positions of points on the wetland polygons to eliminate multiple points collected at one geographic location when there was a pause in navigation or when travel was slow due to obstacles to create smoother polygons. Additional polygons were created in ArcGIS using coordinates provided by PBT and 1998 CIR Digital Orthophotography Quarter Quadrangles (DOQQ) obtained from the North Carolina Department of Transportation. Polygons of all 13 wetlands were combined into one data layer.

The area of each polygon was determined by identifying each vertex along the perimeter of the polygon and its relative contribution to total area as described by Jensen (2005). Table 3 shows the code name and size of each core wetland in the study area with an average wetland size of 0.8111 ha.

Table 3. Name and size of each wetland in the study area.				
Site Code	Area in Hectares	Percentage of Total	# Random Data Points Created	Bog Turtle Population?
Ashe 10	0.3024	2.87%	29	✔
Ashe 15	1.9884	18.86%	189	✔
Ashe 18	3.5943	34.09%	341	✔
Ashe 19	1.3413	12.72%	127	✔
Ashe 5	0.3155	2.99%	30	✔
Ashe 9	0.3911	3.71%	37	✔
Wet BB	0.0320	.30%	3	
Wet C	0.2631	2.5%	25	
Wet D	0.3692	3.5%	35	
Wet DA1	1.0744	10.19%	102	
Wet DA2	0.1835	1.74%	17	
Wet S	0.3247	3.08%	31	
Wet T	0.3646	3.46%	34	
Average Size/Total:	**0.8111**	**100%**	**1000**	

4.1.9 Presence Data

There are a number of inherent biases associated with species occurrence records that must be considered when modeling the spatial distribution of a species. One bias is that researchers most often submit records of occurrence when locating a species in its natural environment, but few reports of species absence are submitted; thus there may be several to a relative abundance of occurrence records, but few absence data records (Peterson and Cohoon 1999; Stockwell and Peters 1999). This scarcity of absence reporting is perhaps a cautious approach, even when an area has been thoroughly searched. Yet, it would not be an acceptable practice for researchers to submit an absence data record for species that are known to be secretive as adaptive survival strategy, physically small and rare in number. Species that fall into this category often require field investigations during different seasons and over long time periods before it can be stated that the species is absent with any certainty. Additionally, an absence record of the species may not be conclusive evidence that it has not occupied the site at some point in the past, and it is possible that the species may migrate to the site at some time in the future if it continues to be suitable habitat.

Araújo and Guisan (2006) noted many presence data records have an inherent bias to roads as this is the easiest and most accessible means to gain access to species' habitats. Stockwell and Peters (1999) and Peterson and Stockwell (2002) cautioned ad hoc museum data are collected using different collection methods and reporting parameters, and the resulting records may lead to different levels of accuracy.

Most bog turtle presence data are associated with a wetland site and only in recent years has GPS accessibility and affordability made it possible to obtain the exact coordinates of individuals within wetland areas. For purposes of this research, it was assumed that if bog turtles were located within a wetland, the turtles could utilize any portion of a wetland, though in reality, there may be some areas that are used and occupied much more than other areas of the wetland and adjacent areas may be frequently or infrequently used as well. Since there have been no studies conducted to date on how bog turtles utilize different regions of each wetland in this research area, Hawth's Analysis Tools' (2004) random script generator was used to generate 1000 presence/absence data points scaled to a percentage of total wetland area for the 13 sites as shown in Table 3 above. Presence data points were placed at the centroid of random 1-meter pixels located within a wetland polygon. Those points created within a wetland polygon known to support a bog turtle population were identified as *presence* data points, and those points located within wetland polygons that are not known to support a bog turtle population were identified as *absence* data points.

4.1.10 Data Processing

All vector data layers were rasterized, resampled and clipped so that all layers were equalized in resolution and spatial extent for the area of interest. Each layer was resampled to 1-meter, 10-meter and 15-meter resolution creating three new raster data layers from each vector layer for analysis in GARP. All other raster imagery and data layers (color and CIR aerial photographs, as well as the DEM and related raster layers) were equalized in spatial extent. Data layers were co-referenced to a standard map projection so that all pixels were in their correct planimetric (*x, y*) locations and projected using North American Datum 1983, Universal Transverse Mercator (UTM) zone 17 north.

4.2 openModeller

The GARP algorithm has experienced a number of implementations since it was created by David Stockwell at Environmental Resources Information Network (ERIN), for the Australian government's Department of Environment, Water Heritage and the Arts. Stockwell implemented the GARP algorithm in a software package and made it available for download via the Internet in 1999; however, GARP was a command-based program that required *a priori* knowledge of computer script-based commands, thus limiting the availability and accessibility of the program to many users. A graphical user interface (GUI) was subsequently developed by Ricardo Scachetti-Pereira in 2003 for the University of Kansas' Natural History Museum Biodiversity Research Center and the Reference Center of Environmental Information (Centro de Referéncia em Informaçáo Ambiental; aka CRIA) in Campinas, Brazil. The GUI made "Desktop GARP," as it was called, more intuitive in its operation and opened its availability to those beyond the sphere of computer programmers.

A new mission emerged during this time at the University of Kansas' Biodiversity Research Center that included mapping the world's biodiversity by using the computational power of Internet users in much the same way as the Search for Extraterrestrial Intelligence (SETI) has been doing at the University of California Berkley. Species occurrence data records were obtained from museums, herbariums and natural history records located all over the world. Those researchers willing to participate in the new "Lifemapper" program could download a screensaver software program that would work in the background to model species distributions in small sequences. These sequences were uploaded to the Biodiversity Research Center's website and another sequence was downloaded and processed (University of Kansas, Natural History Museum, Biodiversity Research Center website and pers. comm. with Renato De Giovanni 06-Jul-2009). Lifemapper now has a new direction and is no longer available in the screensaver format.

Use of the Desktop GARP increased significantly over the next few years and in 2007 the algorithm was modified and implemented in a new GUI developed by Ricardo Scachetti-Pereira while working for openModeller. openModeller is an open-source library of species modeling algorithms that are all included in this newly designed GUI. Most of this work at CRIA was funded by the Foundation for the Support of Research of the State of San Paulo (Fundação de Amparo à Pesquisa do Estado de São Paulo; aka FAPESP) and supported by the Polytechnic School of the University of San Paulo (Escola Politécnica da USP; aka Poli), and the National Institute for Space Research (Instituto Nacional de Pesquisas Espaciais; aka INPE) as an open-source initiative (Sutton et al. 2007)

All models in this research were created using version 1.0.7, GARP with *best subsets - new openModeller implementation* and the *Support Vector Machine* created by Vladimir N. Vapnik and implemented by Renato De Giovanni (Guo et al. 2005). The best subsets is a method of extracting model runs which "fall into the optimal region of the omission/commission graph" as described by Anderson et al. (2003).

Data processing was conducted per the standards set forth in the GARP Users Manual (Stockwell 1999) with 1,000 iterations and maximum convergence of 0.01 as per Raxworthy et al. (2003), McNyset et al. (2005) and Blackburn et al. (2007) and Ortega-Herta and Peterson (2008). This means GARP will continually run selecting rules and statistically analyzing results until either the maximum number of iterations is reached or there is no improvement in predictive accuracy (Blackburn 2007).

The GARP ecological niche model uses one-half of the presence data points to train the model and the other half to verify model accuracy. The GARP model uses a statistical method called bootstrapping where random point data replacement from the original presence data file is used to provide a new collection of presence data points. Data layers representing environmental conditions are used to develop an environmental envelope for areas that are known to support the species of interest. Using a moving window, GARP analyzes each column of variables on a pixel-by-pixel basis to find other areas that fit within the environmental envelope. Each model is built by using a suite of statistical rules, a maximum iterative process and a convergence limit.

4.3 Hardware

This research began with a 32-bit, dual processing computer with a 100 gigabyte hard drive. Initial experiments using openModeller's Desktop GUI took 5 – 6 days using only a few data layers, making it difficult to make even small modifications to parameters or data without waiting several days to note changes in model outcomes. This computational intensity was a condition incurred by the use of high resolution data. A 64-bit computer with 8 processors was made available to run openModeller and the GARP ecological niche model. Only those programs essential to model implementation and review (such as ERDAS, Adobe Reader, ArcMap and Internet connection) were initially loaded on this machine to maximize computational capacity. Additional settings were adjusted in the Windows Task Manager to prioritize computational resources and to tap into the processors to parse the computation load. openModeller was not coded to run multiple threads at this time. Processing time was reduced to approximately 24 hours or less depending on the number of data layers utilized in the model.

4.4 Statistics

All four statistical methods used by the GARP algorithm (atomic, range, negated range, and logistic regression) (Stockwell and Peters 1999) were implemented to produce a final map of areas rated with the highest predicted potential for presence of bog turtle habitats. The simplest of the statistic methods was the atomic rule as it used a single variable value in an "***IF . . . THEN***" rule definition. An example would be that "*IF* silt = .33, *THEN* bog turtle habitats = present." Unlike atomic rules, range rules have a range of values for each data layer. For example, "*if* temperature equals 23 – 29˚C*, then* bog turtle habitats may be present*."* Likewise, the negated range rule was used to reject a range of variable values and omit them from the set of rules.

Rules were customized with "selection, crossover and mutation"

(Sánchez-Flores 2007). Rules are selected using an iterative learning process by analyzing the variable values and developing a tolerance range from all values. Crossover rules combine two ***IF . . . THEN*** rules into one by further refining the tolerance ranges. For example:

Rule 1: ***IF*** silt = .33 AND precipitation = 175 ***THEN*** bog turtle habitat = present

Rule 2: ***IF*** silt = 37 AND elevation = 5200 ***THEN*** bog turtle habitat = present

becomes,

Rule 3: ***IF*** silt = .33 *AND* elevation = 5200 ***THEN*** bog turtle habitat = present

and

Rule 4: ***IF*** silt = .37 *AND* precipitation = 175 ***THEN*** bog turtle habitat = present.

Rule mutation can occur when two conjunctive rules are recombined by GARP to explore every possible combination that will lead to greater predictive ability of the final model.

Logistic regression was used to determine the probability *(P)* of bog turtle habitat occurring at a particular point. If the inclusion of a variable increased model accuracy, the variable was selected. If the overall value of *P* was greater than .75, then the rule was maintained; if the value of *P* was less than .75, the rule was discarded. The rules therefore, became a set of biophysical environmental conditions that define a range of values or tolerances in which the species may be found.

GARP used a combination of these rules internally to produce models by continually improving predictions with an iterative, evolutionary process. It used a bootstrapping method to randomly selected one-half of the species point data set to train the model; the other half of the species data set was used to verify model accuracy. GARP used the bootstrapping technique to determine whether rules were adding to model predictive accuracy and deleted those rules which failed to increase accuracy.

Heuristic rules development was conducted by the model using training data to determine environmental parameters and acceptable ranges. This process may be exceptionally informative as the GARP model may uncover *hidden variables* or synergistic data that haven't been apparent or documented in the past. This would indicate that the combined effect of two or more habitat variables or characteristics are more significant than the sum of their individual effects and it is only when these variables occur simultaneously that specific habitat requirements are met. Descriptive statistics were run on all variables for the presence data points that were created for model implementation. These statistics will provide environmental tolerance ranges for bog turtles.

Jackknifing was used to determine how data layer inclusion or exclusion affected model omission statistics without overfitting the data due to a duplication of environmental variables based on the same data set. Such was the case with the soils data where a large number of variables were created from the same soils database. This was helpful in defining the key variables and relieving the model from the potential to overestimate predictions. A confusion matrix was created for each GARP model and analyzed for accuracy as was the *area under* *the curve* (AUC).

4.5 Mapping Model Results in ArcMap

Fifty models were run using the openModeller GARP algorithm. Out of these fifty models, twenty models with the lowest omission error were selected for analysis in ArcMap. Raster cell values for each data layer were converted to integers and each image was classified with 11 classes representing probability expressed as a percentage of suitable habitat (Table 4). Pixel values in Class 11 with data values ≥ 90 percent were exported for each of the twenty models as a separate data layer and brought back into a new model representing only those potential habitat areas that had a prediction of ≥ 90 percent representing the highest prediction of habitat suitability. Those pixels that overlapped by 80 percent in each data layer (n = 20; 80 percent = 16 data layers) were extracted and exported as a single data layer. Each raster layer was converted to a feature class and all feature class data layers were merged and dissolved to create one layer that covered all areas that were in the upper 90 percent to 100 percent class range of habitat suitability. This feature class was once again rasterized to a 1-meter data layer. A frequency distribution of the probability values for this

Table 4. Classification of Raster Values

Class	% of Probability
Class 1:	No Data
Class 2:	1 – 10
Class 3:	11 – 20
Class 4:	21 – 30
Class 5:	31 – 40
Class 6:	41 – 50
Class 7:	51 – 60
Class 8:	61 – 70
Class 9:	71 – 80
Class 10:	81 – 89
Class 11:	90 - 100

cumulative data layer revealed that most of these pixels fell into class membership areas with ≥ 95% probability of suitable habitat class membership (Figure 12).

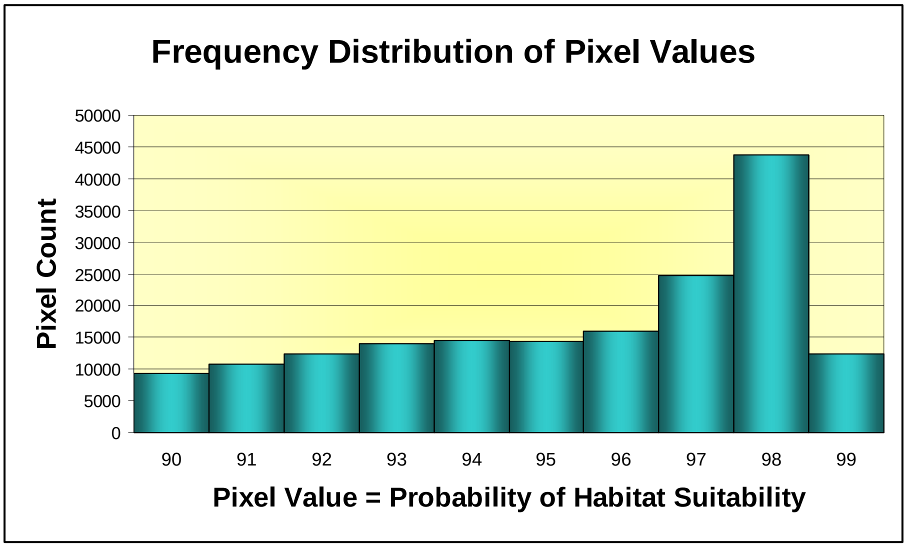

Figure 12. Frequency distribution of the pixel values in the final data layer ≥ 90 percent.

A random grid was designed using the XTools Pro (Data East 2007) for the entire research area (Figure 13). Hawth's Analysis Tools (Beyer 2004) was used to create 1,000 random data points at the centroid of each 1-meter pixel in the final prediction data layer, and within .10 of a mile from a roadway to facilitate ground referencing (Figure 14).

4.6 Ground Referencing

Hawth's Analysis Tools (Beyer 2004) were used to generate points at the centroid of 1,000 random pixels. The *x, y* positions of all 1,000 data points were loaded into a Garmin eTrex Vista HCx hand-held GPS unit for ground referencing purposes. The research area has distinct topographic variations with higher elevations and greater expanses of old-growth forests in the northwest portion of the county and lower elevations and farmlands in the southeastern portion. In an effort to sample from a variety of areas across the county, an attempt was made to visit random data points from different geographic sections of the grid.

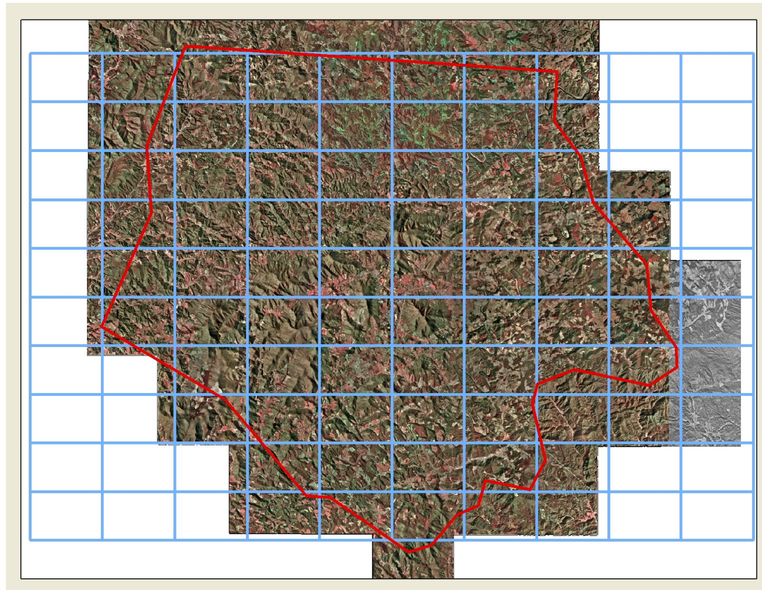

Figure 13. A random grid was designed using the XTools Pro "Create Fishnet" feature for the entire research area.

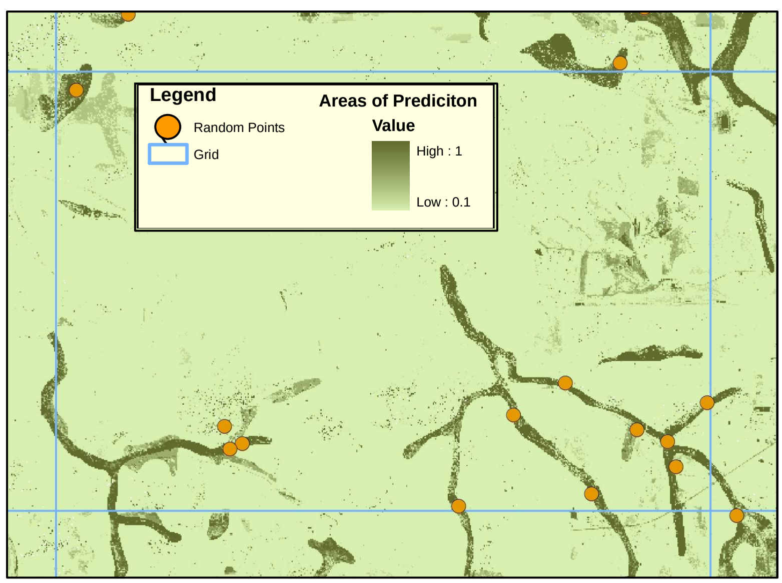

Figure 14. Hawth's Analysis Tools were used to create 1,000 random data points.

When access to the coordinates of a random data point was possible, a white board was used to indicate the coordinates. A field assistant held the white board while pictures were taken so that each picture could be tied back to a specific location. This allowed documentation of the area and a track record of the random data point investigated (Figure 15).

Figure 15. When access to a random data point was possible, a white board was used to indicate the coordinates.

4.7 Confidentiality

Due to the elusive nature of bog turtles, investigation of sites with highest prediction of occurrence will be conducted by experienced biologists from the USFWS, the North Carolina Wildlife Resources Commission (NCWRC) and Project Bog Turtle (PBT). All data that would be considered confidential by one of these agencies will be withheld from public view or discussion in this dissertation.

CHAPTER V

RESULTS

5.0 Overview

This research was designed to evaluate the efficiency and accuracy of using a GIS, remote sensing and the GARP ecological niche model in detecting potential habitat sites suitable to support bog turtles. There have been no known multi-scale studies to date that evaluate these three spatial approaches with multivariate analysis in relation to this species. Due to the high spatial resolution of the data used and number of data layers incorporated into the model, this was an experimental analysis that was restricted to the geographic area of Ashe County, North Carolina. Ashe County has several wetland sites that are known to support bog turtle populations and this county is but a short drive from the University of North Carolina Greensboro.

Previous research by Walton (2006) indicated that high spatial resolution data (>30-meter) are needed to differentiate bog turtle habitat areas that are small in size, non-contiguous and isolated, and that often exhibit heterogeneity due to different land uses and degrees of succession. The current research represents the first known occasion the GARP ecological niche model has been applied to a specialist species using high spatial resolution (1 to 10-meter) data.

Most other models have used 1 km global data sets, even for relatively local or regional applications (Raxworthy et al. 2003; Dominguez-Dominguez et al. 2007; Peterson and Nakazawa 2008; Peterson et al. 2008).

5.1 Data and Model Preparation

This experiment began with a natural color aerial image and a color infrared (CIR) aerial imagery both at 1-meter spatial resolution as it was believed that high spatial resolution would assist in detecting wetland areas. Forty-four soil-related data layers were created with the Soil Data Viewer and SSURGO Soils data; other environmental data layers were originally included as listed in Appendix A. All data layers were projected to Universal Transverse Mercator, 17 North, North American Data 1983; vector data layers were rasterized and resampled to 1-meter resolution and clipped to the same spatial extent as the CIR and natural color imagery.

The original computer used in this research was a 32-bit computer with a dual processor and 75 gigabytes of hard drive storage space to accommodate the data layers used in the model, and to temporarily save and store model results. Initial experiments included the CIR image and eight to ten of the 1-meter raster data layers. The openModeller *GARP with best subsets – DesktopGARP implementation* and the *GARP with best subsets – new openModeller implementation* algorithms (2008) were selected with parameters set to 1,000 iterations and 0.01 convergence level as per Raxworthy (2003), Blackburn (2005) and McNyset (2006). All other parameters were left in their default settings.

However, a number of difficulties were encountered which initially made the possibility of using this methodology and the implementation of the Desktop GARP algorithm in openModeller appear dubious. The experiment continually failed to run despite a number of modifications and various attempts with even a few data layers. A consultation with Tim Sutton, an openModeller programmer with CRIA (pers. comm. 24-Mar-08), led to a number of modifications of the data used in the analysis. First, the size of the data files was larger than what was customarily used with openModeller's suite of algorithms. It was impossible to run the model with a large number of data layers that were also large in file size (≥ 6 gigabytes each). As a compromise, all data layers were resampled to 10-meters, with the exception of the aerial imagery (both natural color and CIR) which remained at 1-meter resolution. Layers derived from the aerial imagery, such as the NDVI and SAVI, remained at the 1-meter resolution as well. All data layers that were resampled to 10-meter resolution were also resampled to 15-meter resolution; however, these 15-meter data layers did not improve model predictions and were subsequently omitted from further analyses. The data layers that were resampled to 10-meter resolution allowed the model to function and yet still maintained the level of spectral integrity required to conduct meaningful analysis.

Secondly, while it was originally believed that categorical data could be included in these analyses (Stockwell and Peters 1999; Stockwell and Peterson 2003), additional studies concluded that the GARP algorithm does not accurately analyze categorical or binary datasets (Elith et al. 2006; McNyset 2006). Specifically, range and negated range rules would often include (or exclude) serial categories that did not support model accuracy. These ranges were included (or excluded in the case of a negated range) because the beginning and ending variables improved model accuracy yet the variable(s) positioned in serial order were included as well, even if they were irrelevant or detrimental to model accuracy. This resulted in the removal of a number of data layers from analysis that were categorical or binary in structure, including the hydrology data set, aspect (based on degrees of aspect) and most of the soils and soils related data layers.

Since soils and hydrology data layers were intuitively critical to model accuracy, other means of extracting soil and hydric parameters were implemented to ensure their representation in subsequent model runs. This included using percentages of sand, silt and clay to determine soil texture and percentage of organic matter content and pH levels to identify soil characteristics. A percentage of slope data layer was created from the DEM to provide a proxy for the hydrology; an aspect data layer using radians was also created from the DEM for analysis.

Thirdly, using the GARP ecological niche model with high spatial resolution data required additional computing capacity, local data storage, and nearly one-half a terabyte of data storage space to store model output images and to analyze model results. In contrast to typical applications of the GARP model, which took a few hours to complete (Tim Sutton, pers. comm. 24-Mar-08), model completion for this research took over sixty-two hours using four data layers on a 32-bit, dual processor machine, resulting in lengthy delays between data manipulation and the time it took to obtain model results.

This research required enhanced computing capacity and subsequently a 64-bit computer with four dual processors and 160 gigabytes of hard drive storage space was purchased to accommodate the intensive computing capacity required of this modeling approach. An additional 300 gigabytes of storage space was made available on a department network server. Additional manipulations were made through the Windows Task Manager to optimize processor usage and prioritize model processing. As a result, central processing unit (CPU) usage was increased from 8—9 percent on the 32-bit computer to 17—29 percent CPU usage on the 64-bit computer with parsing to all eight processors (Figure 16). This new computing

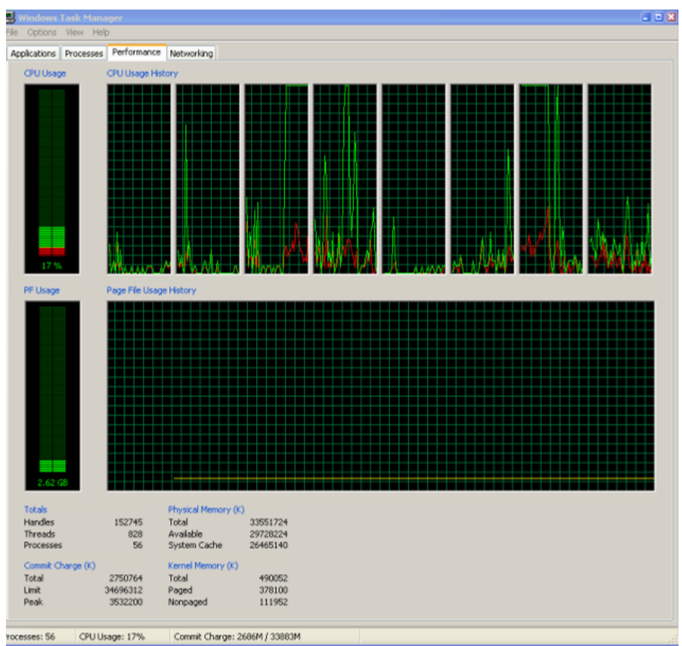

Figure 16. Windows task manager.

capacity increased model efficiency to approximately 1 day or less per run depending upon the number of 1-meter data layers included in each experiment.

5.2 Jackknifing

Many successful applications of the GARP model have indicated that a large number of datasets are required to effectively determine areas of habitat suitability (Raxworthy et al. 2003; Pearson et al. 2006; Blackburn et al. 2005). Other investigations conducted by McNyset (2006 and pers. comm. 30-Oct-2008) indicate that fewer data layers may prevent model overfitting and an over-prediction in areas of data redundancy. Thus, it was more productive to begin with a smaller number of data layers and monitor the response of the omission rate in determining whether specific data layers should be included or omitted based on their contribution to model accuracy. While it was hoped that the inclusion of *more* data layers would help to identify possible hidden variables that are synergistic in nature, this research indicated otherwise based upon the rates of omission.

Omissions are defined as a false negative or the indication that the species is absent in an area when in fact, it is present. Alternatively, commissions, or false positives, are model indications that predict a species is present at a particular location, when if fact it is absent (Loiselle et al. 2003). However, commissions are much harder to quantify, especially if the area is actually prime habitat, but is currently unoccupied or hasn’t been investigated thoroughly to determine species presence/absence status with certainty

(Anderson et al. 2003). Thus omission is a stronger statistic by which to evaluate model accuracy. Lower returns of the omission rate indicate that the model performed with a high degree of accuracy when predicting the other 50 percent (testing data set of the presence data points in areas that are known to support the species of interest.

A jackknife was conducted on all available data layers, as shown in Appendix A, to determine a final list of data layers to be implemented in a trial of fifty models in the GARP ecological niche model. The jackknife process was initiated with the color aerial imagery, CIR aerial imagery, clay, silt, aspect, elevation, NDVI, organic matter and pH. One additional data layer was added and a new model was run to determine how the added layer affected model accuracy. Initial trials to determine which layers were more meaningful resulted in omission errors of approximately 7.0 percent to 13.0 percent. This process continued until the final omission rate was reduced to 0.27397 percent using the eleven data layers as shown in Table 5.

A trial of fifty models was run using the openModeller *GARP with best subsets – DesktopGARP implementation* and the *GARP with best subsets – new openModeller implementation* algorithms (2008). The top twenty models with the lowest omission rates out of the total fifty models were loaded into ArcMap for further spatial analysis.

Table 5. Jackknife Analysis of Data Layers

Data Layer	1	2	3	4	5	6	7	8	9	10	11
CIR Image											
Aspect	JK										
Drainage Basin		JK									
Clay			JK								
Flow Direction				JK							
Elevation					JK						
NDVI						JK					
NDVI-SPOT							JK				
Organic Matter								JK			
Sand									JK		
SAVI										JK	
Silt											JK
Process Time (hours. minutes, seconds):	23.13.14	20.28.33	46.15.58	31.04.39	24.18.34	22.04.15	27.52.08	25.31.12	29.28.55	23.48.49	25.16.07
AUC	**0.97**	**0.99**	**100.00**	**0.98**	**0.97**	**0.99**	**0.99**	**100.00**	**100.00**	**98**	**99**
Accuracy	99.041	97.5342	99.726	98.9041	98.0822	99.589	99.3151	94.6575	99.4521	99.0411	99.1781
Omission	0.95890	2.46575	**0.27397**	1.09589	1.91781	0.41096	.684932	5.34247	0.54795	0.94530	0.82192
% Predicted	11.2647	6.08553	14.1581	11.8896	13.2663	8.43535	8.75207	1.54874	12.39	9.45298	7.23855

JK = Jackknife; this layer was omitted during this model run.

5.3 Final Status of Imagery and Data Layers Used in the GARP Model

5.3.1 Aerial Imagery

The NAIP natural color and CIR imagery were used in the initial executions of the data model. Additionally, the CIR image was used to create the NDVI and SAVI data layers. These data layers were maintained at the original 1-meter resolution to assist in the detection of wetland areas that are small, non-contiguous and experiencing various stages of succession. A jackknifing process indicated that the natural color imagery did not increase model accuracy; however the CIR, NDVI and SAVI were instrumental in model predictions.

The CIR image is a measure of reflectance in the infrared range of the electromagnetic spectrum (700-900 nm); greener, healthier vegetation has a higher reflectance value with possible cell values ranging from 0 to 255. The statistical values for the CIR, NDVI and SAVI are shown in Table 6. NDVI is a measure of photosynthetic activity in an area and values range between -1.0 for barren areas to 1.0 for areas dense in vegetation with high levels of photosynthetic activity. SAVI is a measure of soil reflectance with values ranging between -1.0 for wet soil and 1.0 for exceptionally dry soils or barren rock.

Table 6. Statistics for Aerial Imagery Data Layers

Data Layer:	N	Minimum	Maximum	Mean	Std. Deviation
Ashe Color Infrared Image	730	44	250	136.64	40.418
Normalized Difference Vegetation Index	730	-.573	-.022	-.157	.0667
SAVI	730	-.855	-.0336	-.235	.0996

5.3.2 *Digital Elevation Model Data*

The National Elevation Data (NED) 30-meter Digital Elevation Model (DEM) used in this research indicated that elevation ranges were between 653 meters (2141 feet) at the lowest point in the county to 1588 meters (5210 feet) at the highest point of elevation (Figure 17). The DEM was used to create a drainage basin, flow direction and stream order data in ArcMap. Elevation, drainage basin and flow direction increased model accuracy according to the jackknifing process; the stream order data layer was categorical in nature and could only be used as reference data.

Data analysis revealed that all 13 wetlands in this study (six inhabited by bog turtles and seven that were not) were either in close proximity (≤ 30 meters) or transected by a stream. The six inhabited wetlands were located in four different subbasins in Ashe County along second and third order streams (Strahler 1957). Statistics for the DEM data are shown in Table 7 below.

The 10-meter DEM was used to create an Aspect data layer based upon radians for analysis in the GARP model. In order to convert radians back to degrees, the following formula was used:

$$\textbf{Degrees} = \textbf{radians x } \frac{\mathbf{180}}{\pi}$$

Table 7. Statistics for Digital Elevation Data Layers					
Data Layer:	**N**	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
National Elevation Data 30-meter	730	879.683	959.020	920.710	27.655
Flow Direction	730	1	128	50.49	50.954
Aspect (degrees)	730	287.05	113.39	53.97	52.23
Slope (percentage)	730	1	35	5.54	6.98

The data indicated that the mean direction, in degrees, for the 6 occupied wetlands was 53.97˚.

5.3.3 Weather and Climate Data

Due to the relatively small geographic extent of the research area, certain data layers contained little variability and thus interfered with model accuracy. For instance, there appeared to be distinct divisions in model results as shown in some of the habitat probability maps. The habitat probability map shown in Figure 18 contained three distinct spatial variations: the top portion of the image, Area 1, contained the lowest probability of species habitat suitability; Area 2 suggested medium probability; and Area 3 suggested high probability of finding suitable habitat. The cause of this anomaly was eventually detected by using the jackknife procedure when it was discovered that both precipitation and temperature data layers lacked local variability. These data were based upon three weather stations in Ashe County with interpolated values between weather stations resulting in little variation. Global weather data sets included with the openModeller were implemented with the same results.

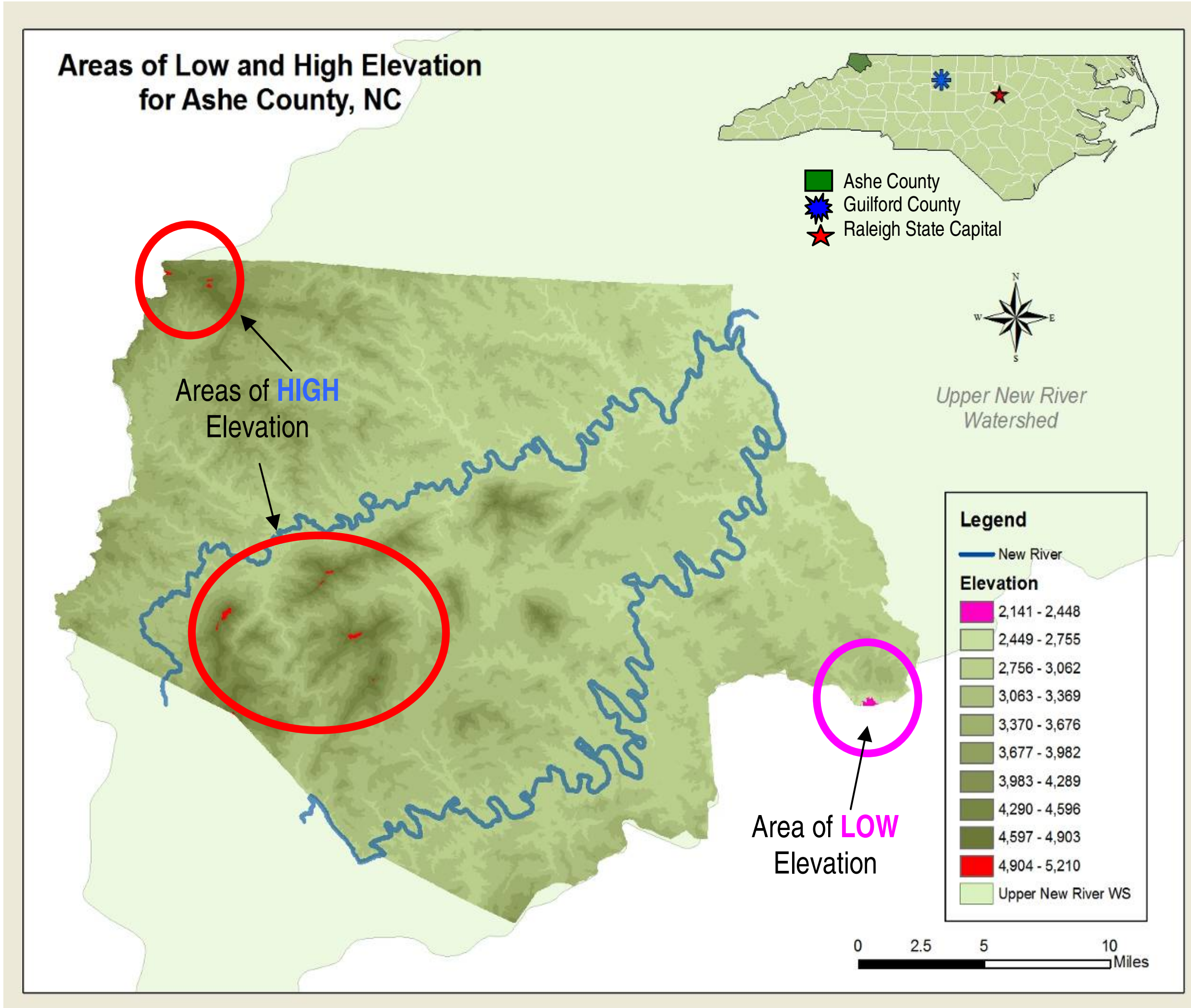

Figure 17. Ashe County elevation ranges were from 653 meters (2141 feet) at the lowest point to 1588 meters (5210 feet) at the highest point of elevation.

PRISM (Parameter-elevation Regressions on Independent Slopes Model) climate data files were obtained from the United States Department of Agriculture's Data Gateway (www.datagateway.nrcs.usda.gov). PRISM weather data is the analysis of point data with a digital elevation model (DEM) that considers weather variations in relation to elevation gradients when interpolating weather values (www.wcc.nrcs.usda.gov/climate/prism). Although there were likely variations in actual local temperatures within the research area of Ashe County, these variations were not captured by the weather data layers; thus model outputs continued to return unreliable data as a result of climatic homogeneity. While other investigations have successfully used weather and climate data as a significant environmental variable in predicting habitat distributions (Peterson and Cahoon 1999: Raxworthy 2003; Chefaoui et al. 2005; Blackburn et al. 2007; Peterson and Nakazawa 2008; Peterson et al 2008), the precipitation and

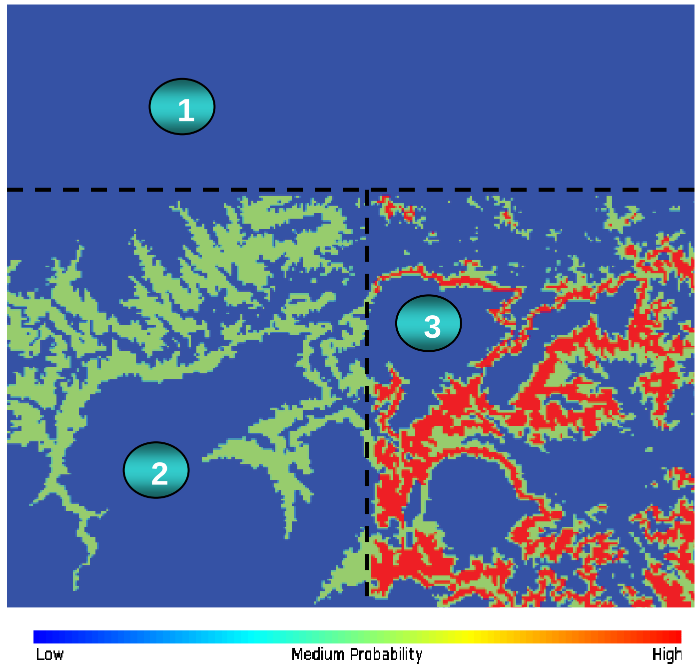

Figure 18. Habitat probability map. Area 1 indicated no probability; Area 2 suggested medium probability; and Area 3 suggested high probability of finding suitable bog turtle habitat.

temperature data were omitted in this investigation due to a lack of spatial variability.

5.3.4 Soils Data

It was fully anticipated that SSURGO soils data and additional data layers created using SSURGO data and the Soil Data Viewer would be instrumental in defining wetland habitats that could support bog turtles. One of the defining characteristics of bog turtle habitats is that they are located in wetland fens containing gleyed, anaerobic soils. However, the soils data used in this model didn’t fully support the hypothesis that soils and/or soil related properties would emerge as critical in identifying additional wetland habitats. Only sand, silt and organic matter improved the omission scores; eliminating the clay data layer brought the omission scores down to 0.27 percent. This, in large part, was due to the fact that these data layers were basically categorical in nature and did not function as continuous data in the GARP modeling process. Of the forty-four data layers created using SSURGO data, only the percentages of sand, silt, clay, and organic matter were used in this research. However, even though model omission rates were reduced with the inclusion of three of these layers (sand, silt and organic matter), it is possible that the GARP model recognized these layers as characteristically categorical data because the frequency distribution for each of these four data layers fell into three frequency classifications as shown in the histograms in Figure 19. For example, if the percentage of sand was 43.5 percent, the other values were exact with no variance in the numeric values: silt was equal to 39.5 percent, clay was equal to 17.0 percent and organic matter was equal to 4.7. If there was a given value for any one of these four data layers, the other three data layers remained consistent in their numerical value for every presence data point without *any* variation.

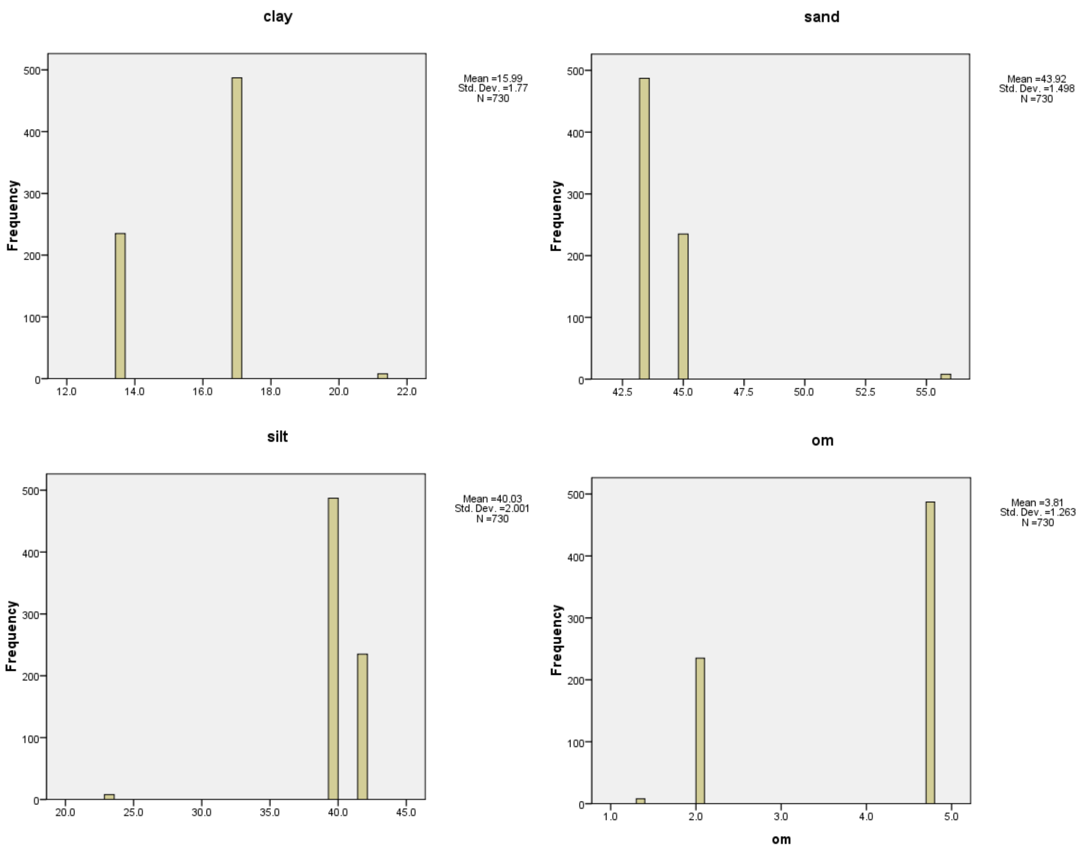

Figure 19. Histograms of the frequency distribution of the clay, sand, silt and organic matter data layers fell into three frequency distribution classes**.**

The problem with this type of data redundancy is not only does it invoke suspicion with respect to the data serving as a proxy to real-world conditions, it may also cause the model to "overfit" the data. Overfitting can result when the model indicates suitable habitat in an area based on a redundancy of data values which leads to an over-prediction of the true distribution of habitat areas. If the categorical soils data were included, the GARP model would suggest suitable habitat could be found in abundance without teasing out specific soil parameters that truly identify the edaphic conditions in these wetland areas. Data redundancy thus prevented the discovery of emergent properties or synergistic qualities that may actually be present in real-world wetland areas, but were not captured by the SSURGO data in a manner that was useful for modeling these layers in the GARP model. These issues of overfitting, data redundancy and jackknifing returns were also reasons that led to the decision to use fewer data layers as opposed to more.

5.3.5 Presence/Absence Data

There were originally 1,000 random data points created for use as presence data in the GARP model. The number of data points assigned to each wetland area was determined by the percentage the wetland's contribution to total area for the thirteen wetlands used in this research. These areas represented six wetlands that were known to currently support a bog turtle population and seven that appeared to be prime habitat, but to date no evidence has been discovered to suggest that bog turtles have historically or currently utilize(d) these areas. The number of data points created per wetland area is shown in Table 3 (Chapter 4, Section 4.1.8).

It was the original intent to code those wetland areas lacking evidence of bog turtle occupancy as a separate species to determine whether or not the GARP model would be able to detect differences between occupied and unoccupied wetlands. However, it was impossible to say with certainty that these apparently unoccupied areas did not indeed have a relict population, or that they haven't been used by bog turtles at some historical point in time, though the population has since dispersed or become extirpated (Chefaoui et al. 2005). It was also possible that these areas were indeed prime habitat that could be used in the future as a natural migratory destination or habitat corridor, as refugia, or as a headstarting location. Thus, in an effort to reduce the potential of introducing false negatives (omission), only the 730 data points representing known areas of occurrence were used in the GARP ecological niche model as it could not be proven conclusively that bog turtles do not or have not occupied the other seven wetland areas (Pearce and Boyce 2006).

One particular advantage of the GARP model is that it automatically establishes absence data points by randomly selecting points from cells other than those occupied by the presence data points, thus creating a combination of presence and absence data points for statistical analysis. This is allows statistical analysis of the areas predicted as suitable habitat and the ability to establish commission values. These absence data points are internal to the GARP algorithm and cannot be used for external analysis. The GARP ecological niche model randomly selects 50 percent of the presence data points (n = 730; 50% = 365) to train the model, and the 50 percent to verify model predictive accuracy.

All data layers that were categorical, binary, or discrete in nature were analyzed, but omitted from the data list used in the GARP model. Since most of the soil types were categorical in nature (see Appendix B), only those layers that represented true continuous data were used. Additional data layers may have been appropriate to use in terms of data type, however, a jackknife process did not show an increase in model accuracy (omission rates) and it was believed that inclusion of these data layers in further analyses could lead to overfitting of model predictions (Peterson et al. 2007).

5.4 Environmental Envelope

The environmental envelope is a description of all the local environmental variables that constitute a species' ecological environment including the tolerance range for each variable within which the species can exist. For instance, this could include precipitation and temperature ranges (high and low values); length solar exposure, or soil composition as defined by percentage of sand, silt, clay, pH level and organic matter. In the GARP ecological niche model, these ranges are determined by evaluating the values of each pixel for every data layer that contains a species presence data point. Descriptive statistics can provide meaningful analysis that will lead researchers to general areas that have similar physical characteristics when searching for the subject species.

This research used 730 presence data points with a variety of data layers that were classified as categorical, binary, discreet and continuous data. A number of these data layers did not qualify for use in the GARP ecological niche model, yet the values of these environmental proxies can be used to describe environmental characteristics and tolerance ranges for the species under investigation. Table 8 is a list of environmental proxies used in this research. Items shaded in grey were used in the final analysis of the GARP ecological niche model when the jackknife analysis indicated that the data layer added to model accuracy. Minimum and Maximum values indicate the range of extremes that the species can tolerate at the local level; final predictions of the GARP ecological niche model identified other areas with these same range of values.

Table 8. Descriptive Statistics of Data Layers

Data Layer:	N	Minimum	Maximum	Mean	Std. Deviation
Ashe Color Infrared Image	730	44	250	136.64	40.418
Available Water Supply 1 - 100 cm	730	14.94	18.14	15.93	1.24
Available Water Supply 0 - 25 cm	730	3.9	4.9	4.553	.2665
Aspect Degrees	730	-72.95	113.39	53.97	52.23
Drainage Basins (density in meters)	730	9320	16690	13332.94	2534.841
CEC-7	730	4.1	14.9	12.578	3.3055
Clay Percentage	730	13.5	21.3	15.987	1.7695
Depth to Water Table July	730	30	201	86.92	80.638
ECEC	730	3.1	11.1	8.578	3.5764
Feature Frost	730	140	150	142.04	1.499
Flow Direction (1, 2, 4, 16, 32, 64, 128)	730	1	128	50.49	50.954

Table 8. Descriptive Statistics of Data Layers					
Data Layer:	N	Minimum	Maximum	Mean	Std. Deviation
Kfactor Rock Free	730	1	5	3.71	1.870
Kfactor Whole	730	2	4	2.65	.933
Ksat (saturated hydraulic conductivity)	730	13.176	28.000	27.808	1.541
Liquid Limit	730	33.2	43.8	40.332	4.9458
North Carolina Land Classification	730	21	90	71.49	19.776
NED 30-meter	730	879.683	959.020	920.710	27.655
Normalized Difference Vegetation Index	730	-.573	-.022	-.157	.0667
Normalized Difference Vegetation Index SPOT	730	-47	2	-25.49	6.687
Organic Matter	730	1.3	4.7	3.810	1.2632
Precipitation Min January	730	3.75	3.75	3.7500	.00000
Precipitation Maximum May	730	5.25	5.25	5.2500	.00000
Precipitation Mean	730	3.209	3.250	3.226	.0158
Precipitation January	730	89	98	94.15	2.638
Precipitation February	730	102	108	105.58	1.672
Precipitation March	730	126	132	129.60	1.403
Precipitation April	730	110	117	114.22	1.940
Precipitation May	730	132	137	133.95	1.783
Precipitation June	730	116	127	122.86	2.960
Precipitation July	730	122	133	127.60	3.393
Precipitation August	730	113	124	118.92	3.240
Precipitation September	730	115	121	118.85	1.686
Precipitation October	730	121	124	121.99	1.351
Precipitation November	730	112	117	114.42	1.205
Precipitation December	730	92	98	95.58	1.672

Table 8. Descriptive Statistics of Data Layers					
Data Layer:	N	Minimum	Maximum	Mean	Std. Deviation
Principle Component (Ashe CIR)	730	85.451	399.312	210.413	54.570
Sand Percentage	730	43.2	55.7	43.916	1.4984
SAVI	730	-.855	-.0336	-.235	.0996
Silt Percentage	730	23	41.5	40.03	2.000
Slope	730	1	35	5.54	6.98
Solar Radiation	730	7.610	5.984	1.617	1.376
Temperature Mean	730	9.817	11.583	10.663	.645
Temperature Max January	730	40	55	49.11	5.579
Temperature Max February	730	55	72	65.70	6.455
Temperature Max March	730	103	122	114.81	7.213
Temperature Max April	730	152	171	164.25	7.319
Temperature Max May	730	194	212	205.49	6.907
Temperature Max July	730	244	262	255.41	6.835
Temperature Max August	730	240	257	251.18	6.620
Temperature Max September	730	209	226	219.47	6.257
Temperature Max October	730	160	177	170.74	6.498
Temperature Max November	730	108	124	117.90	6.016
Temperature Max December	730	59	74	68.47	5.634
Temperature Min January	730	-68	-60	-62.93	3.020
Temperature Min February	730	-57	-48	-51.49	3.383
Temperature Min March	730	-15	-6	-9.10	3.475
Temperature Min April	730	28	36	32.67	2.935
Temperature Min May	730	75	82	79.23	2.601
Temperature Min June	730	114	122	118.59	2.911
Temperature Min July	730	136	144	141.07	3.020
Temperature Min August	730	130	139	135.42	3.333
Temperature Min September	730	97	105	102.07	3.020

Table 8. Descriptive Statistics of Data Layers					
Data Layer:	N	Minimum	Maximum	Mean	Std. Deviation
Temperature Min October	730	37	44	40.60	2.257
Temperature Min November	730	-6	0	-1.96	2.351
Temperature Min December	730	-47	-40	-42.68	2.608
Water Content 15 Bar	730	9.1	19.1	15.819	4.669
Water Content 3rd Bar	730	23.0	29.5	27.336	3.065

5.5 Model Interpretation

There were four principal factors that were taken into consideration when analyzing model results. The first was the percentage of the "area under the curve," (AUC) of the "receiver operating characteristic" (ROC) (Hanley and McNeil 1982) which allows accurate comparison between individual model results. The ROC plot is the "a traditional method for showing the relationship between sensitivity and the false positive rate" (specificity) (Cody and Smith 2006, p 312). AUC values range between 0.00 and 1.00. Values of 0.00 to 0.50 denote that the model has no predictive ability; values between 0.5 and 0.7 are classified as having no to low predictive accuracy; 0.7 – 0.9 represent medium or possible predictive accuracy; and 0.90 to 1.0 represents high predictive accuracy (Pearce and Ferreir 2000; Dominguez-Dominguez 2006). The next factor to consider is model accuracy which is a statistical measurement included in the GARP openModeller program of how faithful the model is to the rule-set used to predict suitable habitat areas.

The third factor is the rate of omission. As stated previously, omission is the model's ability to accurately predict presence data points from the unused portion of the training data set. The lower the omission value, the more accurately the model has predicted the environmental tolerance range as established by the presence data points in the training data set. Since commission values cannot be conclusively verified, the omission rates are a better parameter by which to gauge accurate model predictions.

The final factor to consider in model interpretation is the number of cells predicted as being suitable habitat for the species under consideration. Model outputs are in the form of a raster image and an Adobe .pdf report. The report contains statistical results for each model run, an output map of the image result, definitions of the algorithms used, and a brief analysis of model results. The percentage of cells predicted is not a direct correlation to the raster image. As the jackknife analysis in Table 4 above shows, the percentage of cells predicted ranged from 1.55 percent with the highest omission rate, to 14.16 percent for the lowest omission rate. However, the percentage of cells predicted does not include those cells for which there is no prediction, or a 0.00 percentage of prediction. Prediction ranges go from a low of 0.10 percent up to 1.00 percent. The percentage of cells predicted was best analyzed in ArcMap where the data could be classified and the total number of pixels per class could be evaluated.

5.6 Model Analysis

A series of experiments were run using the openModeller GARP ecological niche model and the various data layers acquired for this research. Overfitting was recognized when the omission rate and the area under the curve (AUC) remained consistent – especially when using several data layers derived from the same data source such as SSURGO data in the models (Table 8).

The jackknife procedure identified data layers whose inclusion in the GARP ecological niche model resulted in lower omission rates. These data layers are highlighted in Table 8 above. The GARP ecological niche model is stochastic in nature and it is impossible to obtain the exact results time after time, even with exact methodology and model execution (Anderson et al. 2003). Thus, a total of 50 models were run out of which the top 20 with the lowest omission errors were used to find areas of common prediction (McNyset 2005; Blackburn et al. 2007). Each model result included an image file that was loaded into a GIS for analysis. Average area under the curve (AUC) was 97 percent for the top 20 models indicating a greater measure of predictive accuracy using the rule-sets created in GARP than random predictions (Pearson 2007). A confusion matrix showed that the model predicted 98.889 percent of the data points accurately (Table 9).

Table 9. Confusion Matrix	**Recorded Present**	**Recorded Absent**
Omission	98.889% (True Positive)	0% (False Positive)
Commission	1.111%(False Negative)	0% (True Negative)

The total research area for this analysis was more than 95,000 hectares. The total number of pixels for the output images of the GARP ecological niche model results was 10,278,323 pixels at 10-meter resolution. The number of pixels that fell into the ≥ 90 percent prediction categories was averaged at 171,829 or 1.67 percent of the total research area predicted as suitable bog turtle habitat. This is a reasonable figure given the endangered species status of the bog turtle and the specific habitat requirements of this species.

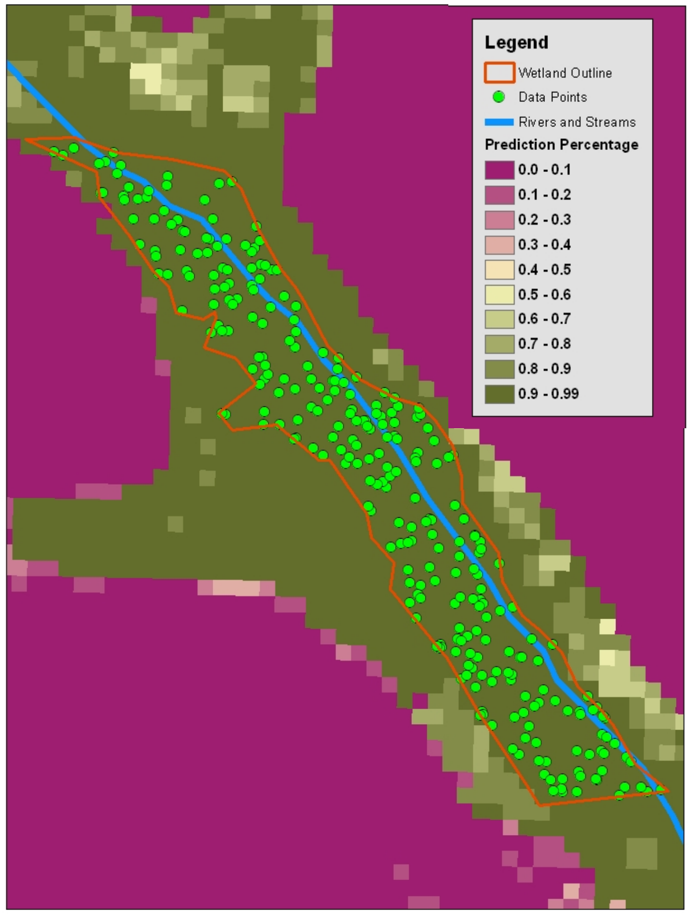

Figure 20. All presence data points projected into environmental space fell into areas of high probability of habitat suitability.

All presence data points (n = 730) projected into environmental space fell into areas of high probability of habitat suitability with a predicted percentage of ≥ 90% (Figure 20). Of particular note, however, was that while the seven unoccupied wetland areas were detected in the high predicted percentage, these areas were at the lower extremes (90 percent to 92 percent) of the prediction range.

5.7 Ground Referencing Model Results

Ground referencing was scheduled for the leaf-off, winter period to see if additional characteristics could be observed that were different from initial observations of the thirteen wetlands in the study area. Overall, ground referencing resulted in 91 data points across the research area. It should be noted that most of the areas of prediction occurred on private property and most landowners are reluctant to allow strangers onto their property without considerable introductions and explanations as to the nature of the visit (*Are there connections between the investigators and government agencies?*), questionable outcomes and perceived threat (*What if a wetland or bog turtles are discovered on my property? Will I lose my land?*), and the potential for future visits. Farmers often cite the fear of introducing livestock diseases, such as hoof-and-mouth disease (*Aphtae epizooticae*) as a reason to refuse access to their property. Many attempts to gain access to interior areas were met with “No Trespassing” and “Private Property” signs. In an effort to reduce the potential for landowner concern, it was decided that no attempts would be made to cross onto private property and that all ground referencing would take place on adjacent roadside areas. This still resulted in a number of instances where paved roads dissolved into dirt roads and progress was stopped and new means of access attempted. It should also be noted that while pictures were taken from the roadside of areas known to support bog turtles, no entry was made onto these sites during any trips made to Ashe County for purposes of this current research.

Each ground referenced data point was positively associated with a wet area as indicated by the presence of water or the observation of hydric soils. Many of these wetland areas, such as the one in Figure 21, showed a distinct signature in terms of hydrophilic vegetation.

Figure 21. Some areas of prediction showed a distinct vegetative signature.

Figure 22. Analysis in ArcMap showed a positive association with a nearby hydrological source.

A buffer analysis in ArcMap showed that all areas of prediction were within ≤ 30-meters of a source of hydrology (Figure 22) as was typical of the wetlands used in this analysis. One area showed distinct signs of bog iron draining off the field which is a common component of bog turtle habitat (Figure 23) (Herman 2003).

In a final analysis, the random point generator was used to find the centroid of each pixel in the raster layer with areas of highest prediction. A kernel density analysis was used to calculate the density magnitude of the data points in the research area to show where the predicted pixels were concentrated. This function produces a contour density map (Figure 24). The density analysis broke the area into 20 classes with a spatial resolution of 10 meters using the GARP model prediction layer as a mask. This information will assist in concentrating search efforts in areas that have a high prediction return from the GARP ecological niche model and a greater concentration of these pixels within a 10 meter radius. Due to the confidential nature of this research, only a small area is shown in Figure 24 for illustrative purposes.

Figure 23. One area showed distinct signs of bog iron draining off the field which is a common component of bog turtle habitat.

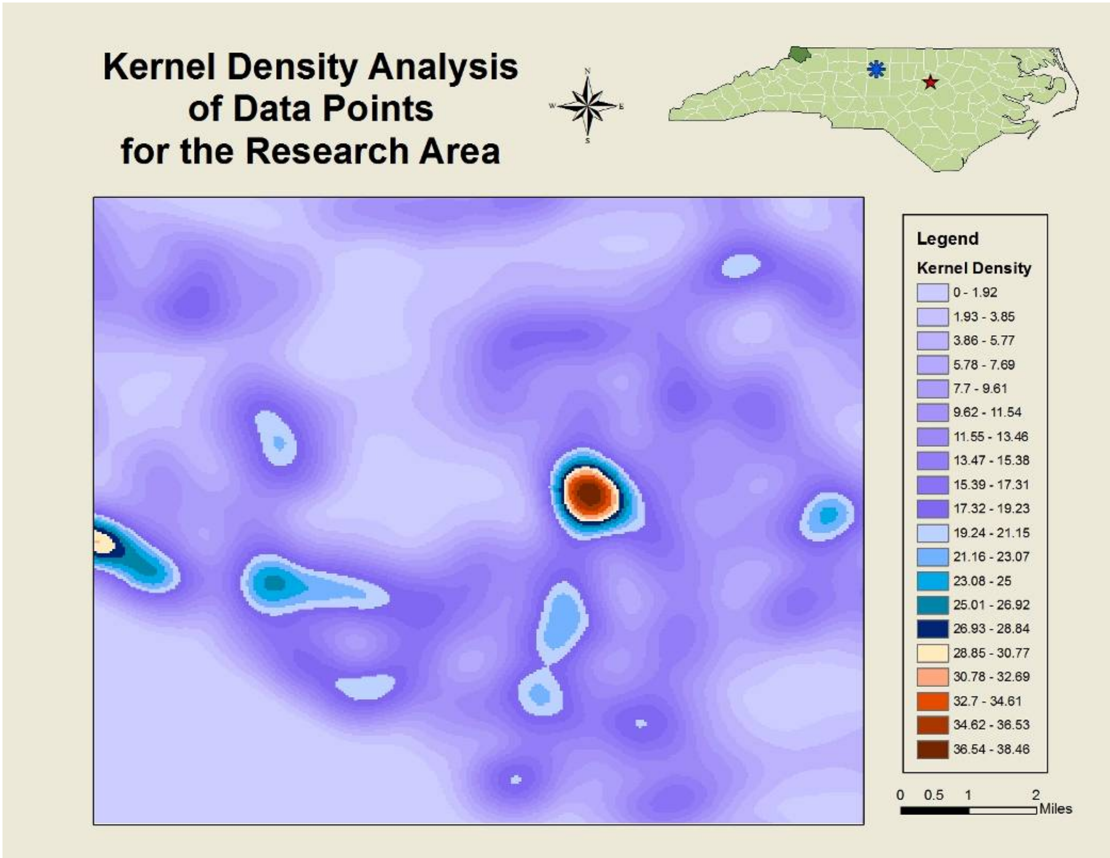

Figure 24. Kernel density analysis finds areas with the highest concentration of data points and produces contours of density on a map.

CHAPTER VI

CONCLUSIONS, DISCUSSION AND FUTURE WORK

This dissertation represents the first known application of the GARP ecological niche model in detecting small, noncontiguous wetland areas that may support bog turtles and thus bog turtle conservation efforts using high spatial resolution data (1—10 meters). It was hypothesized that model results in conjunction with ground reference data would help to identify potential wetland habitat areas and that future field investigations would support model predictions with newly discovered areas that are potential habitat sites or that may serve as migratory corridors, refugia, or head-starting/relocation program locations. While this research is preliminary in terms of covering the entire geographic range for the southern population, the results are promising.

Model results were supported with field investigations to determine the presence of hydric soils, hydrophilic vegetation, bog iron or a combination of these three components. Maps of areas with the highest percentage of habitat suitability have been created along with kernel density estimations to illuminate those areas that have a higher concentration of potential habitat sites. Many of the areas of highest prediction are contiguous in nature, thus it is now possible to infer migratory routes and connecting habitat corridors based on this new data.

The results of this research will be relayed USFWS authorities and the North Carolina Wildlife Resources Commission in the near future who have personnel with field expertise in the location of new wetland habitat areas and the discovery additional bog turtle individuals. These areas can be investigated further to determine whether or not they contain bog turtles. The roles of GISc, GIS, remote sensing and GARP ecological niche modeling were instrumental in conducting these spatial analyses of the environmental proxies to detect and extract areas of similar composition and potential support for bog turtles.

The GARP ecological niche model assisted in the creation of an environmental envelope that established the environmental tolerance ranges for the variables used in this analysis representing a local analysis. Additional information was obtained from other data layers that were categorical in nature and could not be implemented in the model, yet still provide meaningful insight into the overall description of the environmental envelope. This information not only assisted in defining the environmental envelope for the research area of this dissertation, it will also assist in the development of additional models in other areas of the bog turtle's southern range. Understanding the environmental extremes to which this species can endure through descriptive statistics will assist in bog turtle conservation and management strategies. These statistical analyses will also be relayed to the appropriate authorities.

It was hypothesized that soils and soils related data layers would be fundamental in describing the environmental envelope for this species as would data layers that were proxies for hydrology. However, additional soil relationships were not as apparent due largely to the overwhelming number of data layers (91 percent) that had to be omitted from analysis due to the fact that they were categorical or binary in nature. Still, out of the 12 data layers used in the GARP ecological niche model, 5 of the data layers, or 42 percent, were related to soil characteristics (clay, sand, silt, organic matter, and SAVI). River and stream networks for the study area were also omitted as they were treated as binary by the GARP algorithm, yet drainage basin and flow direction were instrumental in model performance.

Additional data layers that were critical in model implementation were the CIR Imagery, the NDVI created from the CIR imagery and the NDVI created from the SPOT image. The spectral reflectance values revealed by these data layers detected areas that aided in the discrimination of wetlands that have either a more constant source, or a longer exposure to hydrologic input. The aspect data layer identified the average angular direction that the wetlands in this study were facing.

The jackknife procedure was instrumental in evaluating the data layers and selecting those data that added to model predictive accuracy without overfitting model results due to data redundancy. While this research was initiated with more than sixty data layers that were available for analysis, the jackknife process assisted in streamlining the available list of variables down to those that were statistically significant through evaluation of the omission rate.

This process alone will save countless hours when investigating other areas for suitable bog turtle habitats. Coordinating this technique with the suite of environmental variables that were used successfully in this research will further aid in reducing the time necessary to locate additional sites that could serve as bog turtle habitat as well as refining the environmental envelope for this species.

This research produced a range of environmental tolerances that describe bog turtle habitats and represented the first analysis of a number of environmental variables and will provide a description of where suitable bog turtle habitats are most likely and least unlikely to exist. However, it should be emphasized that this study only includes a portion of one county within the bog turtle’s southern range. It is possible that there are additional environmental gradients and/or that the environmental tolerance ranges may shift as additional counties are investigated. What is still lacking is an over-arching theory that will explain where bog turtles exist across the southeast and perhaps why some wetlands are occupied while others are not. Future investigations of adjoining counties throughout the southern region should help to shed understanding on these subjects.

One of the inherent problems with presence data is that it is strongly correlated with road networks as this is the easiest means of ingress and egress for researchers investigating new or existing areas. This current research followed this same pattern due to time constraints that are required to gain landowner trust and access to private property. Future research should include an investigation of areas of highest predictions on these privately-owned properties. It has been noted that North Carolina, as well as other areas throughout the bog turtle's range, are rapidly losing historic farmlands and livestock operations (Tesauro and Ehrenfeld 2007). As such, it will be critical to investigate those areas with the greatest potential of bog turtle habitat to implement conservation strategies. It would also be beneficial to obtain an Institutional Animal Care and Use Committee Protocol to allow access onto state-owned or conservation properties to evaluate model predictions.

Future research will include methods to refine the search area even further such as the elimination of flood zone areas associated with the two major forks of the New River in Ashe County. This should reduce the estimation for potential bog turtle habitats considerably though it may mean that more private land areas will need to be investigated. It will also be interesting to investigate further the differences between those wetlands that *do* support bog turtles and those that appear to be suitable habitat, but no turtles have yet been found. Future work will also include an analysis and comparison of the openModeller GARP with *best subsets - new openModeller implementation* algorithm and the *Support Vector Machine* algorithm created by Vladimir N. Vapnik and implemented by Renato De Giovanni (Guo et al. 2005).

Prediction maps have been generated and will be given to trained field researchers from Project Bog Turtle (PBT) and the North Carolina Wildlife Resources Commission (NCWRC) to investigate areas of greater prediction to determine bog turtle presence/absence status. Empirically, this research will be deemed a success if even *one* additional bog turtle (current or historic evidence) is located as a result of these efforts. Theoretically, even if bog turtles are not found in areas of high prediction, it is possible that they may have existed in these areas historically and the identification of suitable habitat will add to the scientific knowledge base as areas that may be used for migratory corridors, refugia, head-starting or relocation programs.

This research presented a unique opportunity to incorporate GISc with conservation efforts. These cutting-edge, state-of-the-art technologies could become the keystone of conservation strategic planning that could propagate new methodologies across many disciplines. Conservation biology, a crisis discipline, may be equipped with modern technologies to address issues of species declines with greater efficacy and positive results. In turn, this information will serve to assist researchers, land managers and policy makers in setting conservation priorities for many endangered chelonian[1] species, including the bog turtle.

[1] Collective term referring to turtles and tortoises; a reptile of the order Chelonia.

REFERENCES

Adjemian, J.C.Z., Girvetz, E.H., Beckett, L., Foley, J.E. 2006. Analysis of genetic algorithm for rule-set prodution (GARP) modeling approach for predicting distributions of fleas implicated as vectors of plague, *Yersinia pestis*, in California. Journal of Medical Entomology 43(1): 93-103.

Akçakaya, H. R. 1996. Linking GIS with Models. A paper presented at the Third International Conference on Integrating GIS and Environmental Modeling, at Santa Fe, New Mexico, on January 23, 1996.

Altherr, S. and D. Freyer. 2000. Asian turtles are threatened by extinction. Turtle and Tortoise Newsletter, 1:7-11.

Amon, J. P., C. A. Thompson, Q. J. Carpenter, and J. Miner. 2002. Temperate zone fens of the glaciated midwestern USA. *Wetlands,* 22:301–317.

Anderson, R. P., D. Lew, and A.T. Peterson. 2003. Evaluating predictive models of species' distributions: Criteria for selecting optimal models. *Ecological Modelling* 162: 211- 232.

Araújo, M.B. and Guisan, A. 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography* 33: 1677-1688.

Ashe County GIS tax parcel website (http://ashegis.ashecountygov.com/webgis/)

Aspinall, R. J. 1995. Geographic information systems: their use for environmental management and nature conservation. *Parks* 5, p. 20–31.

Beans, B.E. and L. Niles. 2003. *Endangered and Threatened Wildlife of New Jersey.* Rutgers University Press. 336 pp.

Bedford, B. L. and K. S. Godwin. 2003. Fens of the United States: Distribution, characteristics and scientific versus legal isolation. *Wetlands*, Vol. 23, No. 3, p. 608–629.

Behler, J. 1997. Troubled times for turtles. In Van Abbema (Ed.). Proc.: Conserv., Restor. and Manag. of Tortoises and Turtles – An International Conference. July 1993. Purchase, New York: Turtle Recovery Program and New York Turtle and Tortoise Society, pp. xvii-xxii.

Beyer, H. L. 2004. Hawth's Analysis Tools for ArcGIS. Available at: http://www.spatialecology.com/htools.

Blackburn, J.K., McNyset, K.M., Hugh-Jones, M.E., Curtis, A.C. 2007. Modeling the geographic distribution of Bacillus anthracis, the causative agent of anthrax disease, for the contiguous United States using predictive ecological niche modeling. American Journal of Tropical Medicine (forthcoming)

Bond J.E., D.A. Beamer, T. Lamb, & M.C. Hedin. 2006. Combining genetic and geospatial analyses to infer population extinction in mygalomorph spiders endemic to the Los Angeles region. Animal Conservation, 9: 145-157, doi:10.1111/j.1469-1795.2006.00024.x.

Brady, N. C. and R. R. Weil. 2002. *The Nature and Properties of Soil. 13th Edition*. Prentice Hall, Upper Saddle River, NJ.

Braunisch V., K. Bullmann, R. F. Graf and A. H. Hirzel. 2008. Living on the edge—Modelling habitat suitability for species at the edge of their fundamental niche. Ecological Modelling. Volume 214, 2-4, pp 153-167.

Buhlmann, J. A., M. G. Rollins, S. M. Roble, and J. C. Mitchell. 1997. New approaches for the conservation of bog turtles, *Clemmys muhlenbergii*, in Virginia. *In:* J.V. Abbena (ed.) Proceedings: *Conservation, Restoration, and Management of Tortoises and Turtles—An International Conference.* SUNY, Purchase, NY.

Carter, S.L., C.A. Haas, and J.C. Mitchell. 2000. Movements and activity of bog turtles (*Clemmys muhlenbergii*) in southwestern Virginia. J. Herpetol. 34(1):75-80.

Chang, K. 2008. *Introduction to Geographic Information Systems 4th Edition*. New York, NY. McGraw-Hill Companies, 450 pages.

Chase, J.D., K.R. Dixson, J.E. Gates, D. Jacobs, and G.J. Taylor. 1989. Habitat characteristics, population size, and home range of the bog turtle, *Clemmys muhlenbergii*, in Maryland. J. Herpetology 23(4):356-362.

Chase, J.M. & Leibold, M.A. 2003. *Ecological niches – linking classical and contemporary approaches*. The University of Chicago Press, Chicago, IL.

Centro de Referência em Informação Ambiental (CRIA) (Reference Center of Environmental Information) located at: http://www.cria.org.br/.

Chen, G. J., and A. T. Peterson. 2000. A new technique for predicting distribution of terrestrial vertebrates using inferential modeling. *Zool. Res*. 21, 231–237.

Chefaoui, R., J. Hortal and J. M. Lobo. 2005. Potential distribution modelling, niche characterization and conservation status assessment using GIS tools: a case study of *Iberian Copris* species. *Biological Conservation* 122 327–338.

Cowardin, L.M, V. Carter, F.C. Golet, and E.T. LaRoe. 1979. Classification of wetlands and deepwater habitats of the United States. Department of the Interior. U.S. Fish and Wildlife Service, Washington, D.C. 131 p.

Dahl, T.E. 2000. Status and trends of wetlands in the conterminous United States 1986 to 1997. U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C. 82 pp.

Dahl, T.E. and G. J.Allord. “Technical Aspects of Wetlands History of Wetlands in the Conterminous United States” National Water Summary on Wetland Resources, United States Geological Survey Water Supply 2425. Retrieved December 15, 2008 from: http://water.usgs.gov/nwsum/WSP2425/history.html.

Darwin, C. and E. Mayr. 1859. *On the Origin of Species by Means of Natural Selection.* Contributor Ernst Mayr. 1964. *A Facsimile of the First Edition.*. Harvard University Press, 513 p.

Data East. 2007. Developer of XTools Pro version 5.1.0. Available at www.xtoolspro.com.

Davis, F.W., D.M. Stoms, J.E. Estes, J. Scepan and J. M. Scott. 1990. An information systems approach to the preservation of biological diversity. *International Journal of Geographic Information Systems* 4, pp. 55–78.

Domínguez-Domínguez O., E. Martínez-Meyer, L. Zambrano, G. Pérez-Ponce de Léon. 2006. Using Ecological-Niche Modeling as a Conservation Tool for Freshwater Species: Live-Bearing Fishes in Central Mexico. *Conservation Biology* V:20, NO: 6, pg. 1730-1739

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettman, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J., Williams, S.E., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve prediction of species’ distributions from occurrence data. *Ecography* 29, 129–151.

Elton, C. S. 1927. *Animal Ecology.* Sedgwick and Jackson, London.

Engler, R., Guisan, A., Rechsteiner, L. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. Journal of Applied Ecology 41: 263-274.

Ernst, C. H., J. E. Lovich, and R. W. Barbour. 1994. *Turtles of the United States and Canada*. Smithsonian Inst. Press, Washington, DC. 578 p.

ESRI Data & Maps [CD-ROM]. (2008). Redlands, CA: Environmental Systems Research Institute.

Frederick, C. M. 1974. A natural history survey of the vascular flora of Cedar Bog, Champaign County, Ohio. *Ohio Journal of Science* 74:65–116.

GARP Users Manual and Technical Manual located at: http://www.lifemapper.org/desktopgarp/Default. asp?Item=3&Lang=1

Gibbons, J. W., D. E. Scott, T. J. Ryan, K. A. Buhlmann, T. D. Tuberville, B. S. Metts, J. L. Greene, T. Mills, Y. Leiden, S. Poppy, and C. T. Winne. 2000. The global decline of reptiles, déjà vu amphibians. *BioScience* Vol. 50; No. 8: 653- 666.

Grinnell, J. 1917. The niche-relationships of the California Thrasher. *Auk*, 34, 427–433.

Guo, Q. H.; Kelly, M.; Graham, C. H. 2005. Support vector machines for predicting distribution of sudden oak death in California. Ecological Modelling. 182: 75-90.

Hanley, J., and B. McNeil. 1982. The meaning and use of the area under an ROC curve. Radiology 143(1):29-36, 1982.

Hanski, I. and D. Simberloff. 1997. The metapopulation approach, its history, conceptual domain, and application to conservation. *In*: Hanski, I. and Gilpin, M. (eds), *Metapopulation biology. Ecology, genetics, and evolution*. Academic Press, p. 5-/26.

Hefner, J.M., B.O. Wilen, T.E. Dahl, and W.E. Frayer. 1994. Southeast Wetlands: Status and Trends, mid-1970s to mid-1980's. A cooperative publication by the United States Department of the Interior Fish and Wildlife Service, Southeast Region, Atlanta, Georgia and the United States Environmental Protection Agency, Region IV, Atlanta, Georgia.

Herman, D.W. 1989. Tracking the rare bog turtle. Wildlife N.C. 53(10):17-19.

Herman, D. W. 1994. The bog turtle, *Clemmys muhlenbergii,* in North Carolina. An action plan for its conservation and management. Report submitted to the North Carolina Resources Commission. 156 p.

Herman, D. W. 1999. The impacts of livestock grazing on bog turtle habitat in the piedmont and mountains of the Southeast. Report submitted to US Natural Resources Conservation Service, Wetlands Science Institute.

Herman, D. W. and B. W. Tryon. 1997. Land use, development, and natural succession and their effects on bog turtle habitat in the southeastern United States. p. 364-371. *In* J. Van Abbema (ed.), Proceedings: Conservation, Restoration, and Management of Tortoises and Turtles—An International Conference. July 1993, State University of New York, Purchase. New York Turtle and Tortoise Society, New York.

Herman, D. W. 2003. Status of the bog turtle, *Clemmys muhlenbergii,* Schoepff, in the Southern United States. Project Bog Turtle’s final report to the U.S. Fish and Wildlife Service on the 1996—2002 status survey conducted under Grant Agreement #1448-0004-96-9126.

Huber, B. Quantitative Decisions. (http://www.quantdec.com)

Hutchinson, G. E. 1957. Concluding remarks, *Cold Spring Harbor Sump. On Quantitative Biol. 22, Population Studies: Animal Ecology and Demography.* Cold Spring Hrbor, Biol. Lab., P. 415 – 427.

Jensen, J. R. 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. Upper Saddle River, NJ: Prentice Hall, 3rd Ed., 526 pages.

Kluza, D.A. and McNyset, K.M. 2005. Ecological niche modeling of aquatic invasion species. Aquatic Invaders 16(1): 1-7.

Lee, D.S. and A.W. Norden. 1994. The distribution, ecology, and conservation needs of bog turtles, with special emphasis on Maryland. Maryland Nat. 40:1-46.

Leica Geosystems Geospatial Imaging Releases ERDAS IMAGINE(R) V9.0

Levins, R. 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America*, 15: 237-240.

Loiselle, B. A., Howell, C. A., Graham, C. H., Goerck, J. M., Brooks, T., Smith, K. G. & Williams, P. H. 2003. Avoiding pitfalls of using species distribution models in conservation planning. *Conservation Biology* 17: 1591-1600.

MacArthur, R. J. and E. O. Wilson. 1967. *The theory of island biogeography*. Princeton University Press, Princeton, NJ.

McNyset, K.M. 2005. Use of ecological niche modeling to predict distributions of freshwater fish species in Kansas. *Ecology of Freshwater Fish* 14 (3): 243-255.

Mitsch, W.J. and J.G. Gosselink. 2000. The value of wetlands: importance of scale and landscape setting. *Ecological Economics* 35: 25-33.

Munsell Color. 1990. Munsell soil color charts. Munsell Color, 2441 North Calvert Street, Baltimore, MD 21218.

Natural Resources Conservation Service. 2002. *Habitat Notes: Meadow Bogs.*

National Oceanic & Atmospheric Administration. Uncovering Chameleons. A case study based upon the work of Raxworthy et al (2003). Located at: http://earthobservatory.nasa.gov/Study/Chameleons/chameleons3.html.

National Oceanic & Atmospheric Administration. PRISM data obtained from: (ftp://ftp.ncdc.noaa.gov/pub/data/ prism100/).

Neerinckx, S., A.T .Peterson, H. Gulinck, J. Deckers and H. Leirs. 2008. Geographic distribution and ecological niche of plague in sub-Saharan Africa. *International Journal of Health Geographics* 7:54.

Nemuras, K.T. and J.A. Weaver. 1974. The bog turtle: synonym for extinction? Nat. parks Conserv. Mag. 48(6):17-20.

North Carolina Department of Transportation http://www.ncdot.org/it/gis/.

North Carolina Drought Management Advisory Council located at: www.ncdrought.org/.

North Carolina General Assembly website: www.ncga.state.nc.us/.

North Carolina State Climate Office located at www.nc-climate.ncsu.edu.

North Carolina Wildlife Resources Commission. 2006. Bog Turtle Fact Sheet. 1722 Mail Service Center Raleigh, N.C. 27699-1722. (www.ncwildlife.org)

openModeller. openModeller is a fundamental niche modelling library, providing a uniform method for modelling distribution patterns using a variety of modelling algorithms. http://openmodeller.sourceforge.net/

Ortega-Huerta, M.A. and A. T. Peterson. 2008. Modeling ecological niches and predicting geographic distributions: a test of six presence-only methods. *Revista Mexicana de Biodiversidad.* 79:205-210.

Pearce, J. L. and M. S. Boyce. 2006. Modelling distrubituion and abundance with presence-only data. *Journal of Applied Ecology.* Vol 43, 405 – 412.

Pearce, J., and S. Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* 133**:**225–245.

Pearson, R.G., W. Thuiller, M.B. Araújo, E. Martinez-Meyer, L. Brotons, C. McClean, L. Miles, P. Segurado, T.P. Dawson, and D.C. Lees. 2006. Model-based uncertainty in species range prediction. *Journal of Biogeography* 33:1704-1711.

Pearson, R.G., C.J. Raxworthy, M. Nakamura and A.T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar *Journal of Biogeography* 34:102-117.

Pearson, R.G. 2007. Species' Distribution Modeling for Conservation Educators and Practitioners. Synthesis. American Museum of Natural History. Available at http://ncep.amnh.org.

Peterson, A.T. 2001. Predicting species' geographic distributions based on ecological niche modeling. Condor 103: 599-605.

Peterson, A. T., and K. C. Cohoon. 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. Ecological Modeling 117: 159-164.Peterson et al. 2002;

Peterson, A.T. and D.A. Vieglais. 2001. Predicting Species Invasions Using Ecological Niche Modeling: New Approaches from Bioinformatics Attack a Pressing Problem. BioScience. Vol. 51 No. 5 p. 363–371.

Peterson, A. T., D. R. B. Stockwell, and D. A. Kluza. 2002. Distributional prediction based on ecological niche modeling of primary occurrence data. *In*: Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A., Samson, F.B. (Eds.), *Predicting Species Occurrences: Issues of Accuracy and Scale*. Island Press, Washington, DC, p. 617–623.

Peterson, A.T., Sanchez-Cordero, V., Beard, C.B., and Ramsey, J.M. 2002. Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. Emerging Infectious Diseases 8(7): 662-667.

Peterson, A. T., M. Papes, and D. A. Kluza. 2003. Predicting the potential invasive distributions of four alien plant species in North America. *Weed Science* 51: 863-868.

Peterson, A.T., Bauer, J.T., Mills, and J.N. Mills. 2004. Ecologic and geographic distribution of Filovirus disease. Emerging Infectious Diseases 10(1): 40-47.

Peterson, A.T. and D.A. Kluza. 2005. Ecological Niche Modeling as a New Paradigm for Large-Scale Investigations of Diversity and Distribution of Birds. USDA Forest Service Gen. Tech. Rep. PSW-GTR-191.

Peterson, A.T., M. Papes¸ and M. Eaton. 2007. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. Ecography 30: 550-560.

Peterson, A. T., M. Papes and Jorge Soberón. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological modelling* V2:13, pg. 63–72.

Peterson A.T. and Y. Nakazawa. 2008. Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri. Global Ecology and Biogeography, (Global Ecol. Biogeogr.)* (2008) 17, 135–144.

Project Bog Turtle. http://www.projectbogturtle.org/.

Primack, R. B. 1995. *A Primer of Conservation Biology*. Sinauer Associates, Sunderland, Massachusetts.

Ramanujan, K. 2004. New Tools for Conservation. NASA website located at: http://earthobservatory.nasa.gov/Study/Conservation/.

Raxworthy, C.J., Martinez-Meyer, E., Horning, N., Nussbaum, R.A., Schneider, G.E., Ortega-Huerta, M.A., Townsend Peterson, A. 2003. Predicting distributions of known and unknown reptile species in Madagascar. Nature 426: 837-841.

Rhodin, A. G. J. 2000. Publisher's Editorial: Turtle Survival Crisis. Turtle and Tortoise Newsletter, 1:2-3.

Rushton, S. P., S. J. Ormerod and G. Kerby. 2004. New paradigms for modeling species distributions? *Journal of Applied Ecology*. Vol: 41, 193–200.

Salzberg, A. 1998. The Chinese turtle problem. *Reptilia*.

Schafale, M.P. and A.S. Weakley. 1990. Classification of natural communities in North Carolina, third approximation. N.C. Natural Heritage Program report. 325 pp.

Scott, D. E. 2005. *Ambystoma opacum*, Marbled salamander. Pp. 627-632 *In:* Amphibian Declines: the Conservation Status of United States Species. M. J. Lannoo (ed), University of California Press, Berkeley, CA.

Scott, J. M., T. H. Tear, and F. W. Davis. 1996. Gap Analysis: a landscape approach to biodiversity planning. *American Society for Photogrammetry and Remote Sensing*, Bethesda, MD.

Search for Extraterrestrial Intelligence (SETI) located at: http://www.seti.org/seti.

Shlens, J. 2009. A Tutorial on Principal Component Analysis. Version 3.01. Center for Neural Science, New York University, New York City, NY 10003-6603 and Systems Neurobiology Laboratory, Salk Insitute for Biological Studies La Jolla, CA 92037

Soberón, J. and A. T. Peterson. 2005. Interpretations of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2, pp. 1-10.

Stockman Amy K., David A. Beamer and Jason E. Bond. 2006. An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. *Diversity and Distributions,* 12, 81–89.

Stockwell, David. Original creator of Genetic Algorithm for Rule-set Production (GARP) at ERIN Unit of Environment Australia and enhanced at the San Diego Supercomputer Center.

Stockwell, D. and Peters, D. 1999. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* 13(2): 143-158.

Stockwell, D.R.B. and Peterson, A.T. 2002. Effects of sample size on accuracy of species distribution models. *Ecological Modelling* 148: 1-13.

Stockwell, D.R.B. and A.T. Peterson. 2003. Comparison of resolution of methods used in mapping biodiversity patterns from point occurrence data. *Ecological Indicators* 3:213–221.

Strahler, A. N., 1952. Dynamic Basis of Geomorphology. *Geological Society of America Bulletin* 63:923-938.

Sutton, T., R. de Giovanni and M. Ferreira de Siqueira. 2007. Introducing openModeller: A fundamental niche modelling framework. OSGeo Journal Vol. 1, May 2007

Tesauro, J. and D. Ehrenfeld. 2007. The effects of livestock grazing on the bog turtle [*Glyptemys (Clemmys) muhlenbergii*]. *Herpetologica*, 63(3), 293–300.

Turner, W., S. Spector, N. Gardner, M. Fladeland, E. Sterling and M. Steininger. 2003. Remote sensing for biodiversity science and conservation. *TRENDS in Ecology and Evolution* Vol.18 No.6.

United States Census Data. Retrieved on Jan 18th from: http://censtats.census.gov/

United Statese Department of Agriculture. 1975. Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys. U.S. Department of Agriculture. Soil Conservation Service, Soil Survey Staff, Agricultural Handbook 436, Washington, D.C.871 pp.

United States Department of Agriculture, Natural Resources Conservation Service, Soil Survey Staff. 2006. *Keys to Soil Taxonomy*, *10th Edition.* U.S. Government Printing Office Washington, DC 20402. 341 pp.

United States Department of Agriculture, Soil Data Mart website located at: http://soildatamart.nrcs.usda.gov/.

United States Department of Agriculture, Natural Resources Conservation Service. Soil Data Viewer located at: http://www.itc.nrcs.usda.gov/soildataviewer/.

United States Department of Agriculture Natural Resources Conservation Soil Data Viewer 5.2 User Guide. September 2007.

United States Department of Agriculture National Agriculture Imagery Program located at: http://165.221.201.14/NAIP.html.

United States Department of Agriculture Natural Resources Conservation

Service Geospatial Data Gateway located at
http://datagateway.nrcs.usda.gov/.

United States Fish and Wildlife Service. 1973. Endangered Species Act and Amendments. Entire document can be found at:
http://endangered.fws.gov/ esa.html.

United States Fish and Wildlife Service. 1997. Final rule to list the northern population of the bog turtle as threatened and the southern population as threatened due to similarity of appearance. Federal Register 62 (213): 59605-59623.

United States Geological Survey Digital Elevation Model located at:
http://ned.usgs.gov/

United States Geological Survey Topographic Maps.

University of Kansas Natural History Museum & Biodiversity Research Center. Website: http://www.nhm.ku.edu/.

van Dijk, P.P. 2000. The status of turtles in Asia. *Chelonian Research Monographs* 2: 15-23.

van Dijk, P. P., Stuart, B., and Rhodin, G. J. (eds.) 2000. Asian Turtle Trade: Proceedings of a workshop on conservation and trade of freshwater turtles and tortoises in Asia. *Chelonian Research Monographs* 2: 164 pp.

Venkataraman, A. B., N. V. Kumar, S. Varma and R. Sukumar. 2002. Conservation of a flagship species: Prioritizing Asian Elephant (*Elephas maximus*) conservation units in southern India. *Current Science*, Vol. 82, No. 8.

Venkataraman, A. R., R. Saandeep, N. Baskaran, M. Roy, A. Madhivanan, and R. Sukumar. 2005. Using satellite telemetry to mitigate elephant–human conflict: an experiment in northern West Bengal, India. *Current Science* 88:1827–1831.

Vogiatzakis, I. N. 2003. GIS-based Modeling and Ecology: A Review of Tools and Methods. Geographical Paper No. 170. Department of Geography, The University of Reading, Reading.

Walton, E.M. 2002. Field Techniques for Bog Turtle Research in North Carolina. E.M. Walton, Greensboro, NC. 68 p.

Walton, E.M. 2006. Using Remote Sensing and Geographical Information Science to Predict and Delineate Critical Habitat for the Bog Turtle, *Glyptemys muhlenbergii.* M.A. Thesis, University of North Carolina Greensboro.

Weaver, J.C. 2005. The drought of 1998–2002 in North Carolina—Precipitation and hydrologic conditions: U.S. Geological Survey Scientific Investigations Report 2005–5053, 88 p.

Wilson, D.S., C.R. Tracy, and C.R. Tracy. 2003. Estimating age of turtles from growth rings: A critical evaluation of the technique. *Herpetologica,* 59(2), 2003, 178–194.

Withgott, J. and S. Brennan. 2009. *Essential Environment: The science behind the stories. 3rd Edition*. Pearson Benjamin Cummings, San Francisco, CA. 456 pp.

APPENDIX A

Data Layers Used in this Research

Data Layer	Source
Ashe Color Infrared Image	§
Ashe CIR PCA	§
Aspect Degrees	•
Available Water Capacity	†
Available Water Supply 0 - 25 cm	†
Available Water Supply 1 - 100 cm	†
Bulk Density 1/3 Bar	†
Bulk Density 15 Bar	†
Calcium Carbonate	†
Cation Exchange Capacity-7	†
Clay Percentage	†
Depth to Water Table December	†
Depth to Soil Restrictions	†
Depth to Water Table	†
Depth to Water Table July	†
Drainage Basins (density in meters)	•
Drainage Class	†
Effective Cation Exchange Capacity	†
Electrical Conductivity	†
Farmland Classification	†
Feature Frost	†
Flooding Frequency Class	†
Flow Direction (1, 2, 4, 16, 32, 64, 128)	•
Frost Action	†
Frost Free Days	†
Hydric Rating	†
Hydrologic Group	†
Kfactor Rock Free	†
Kfactor Whole	†
Ksat (saturated hydraulic conductivity)	†
Liquid Limit	†

Data Layer	Source
NED 30-meter	●
Normalized Difference Vegetation Index	§
Normalized Difference Vegetation Index SPOT	*
North Carolina Land Classification	□
Organic Matter Content	†
Permeability Class	†
pH Rating	†
Plastic Index	†
Ponding Frequency Class	†
Precipitation April	▲
Precipitation August	▲
Precipitation December	▲
Precipitation February	▲
Precipitation January	▲
Precipitation July	▲
Precipitation June	▲
Precipitation March	▲
Precipitation Maximum May	▲
Precipitation May	▲
Precipitation Mean	▲
Precipitation Min January	▲
Precipitation November	▲
Precipitation October	▲
Precipitation September	▲
Principle Component (Ashe CIR)	§
Sand Percentage	†
Saturated Hydraulic Conductivity	†
SAVI	§
Silt Percentage	†
Slope (percentage)	●
Sodium Absorption Ratio	†
Soil Texture	†
Solar Radiation	●
Stream order	●
Surface Texture	†

Data Layer	Source
Temperature Max April	▲
Temperature Max August	▲
Temperature Max December	▲
Temperature Max February	▲
Temperature Max January	▲
Temperature Max July	▲
Temperature Max March	▲
Temperature Max May	▲
Temperature Max November	▲
Temperature Max October	▲
Temperature Max September	▲
Temperature Mean	▲
Temperature Min April	▲
Temperature Min August	▲
Temperature Min December	▲
Temperature Min February	▲
Temperature Min January	▲
Temperature Min July	▲
Temperature Min June	▲
Temperature Min March	▲
Temperature Min May	▲
Temperature Min November	▲
Temperature Min October	▲
Temperature Min September	▲
T-Factor	†
Water Content 15th Bar	†
Water Content 3rd Bar	†

§ **National Agriculture Imagery Program—**U.S. Department of Agriculture, Service Center Agencies, National Cartography & Geospatial Center.

† **Soil Survey Geographic (SSURGO)—**database for Ashe County, NC; U.S. Department of Agriculture, Natural Resources Conservation Service.

● **National Elevation Dataset 10-meter—**U.S. Department of Agriculture Natural Resources Conservation Services, National Cartography & Geospatial Center

* **SPOT Panchromatic Imagery—**(10-meter); obtained from UNCG data library.

□ **North Carolina Land Classification Data—**U.S. Department of Agriculture, National Agricultural Statistics Service, Research and Development Division, Geospatial Information Branch, Area Frame Section

▲ **Precipitation and Temperature Data—**U.S. Department of Agriculture Natural Resources Conservation Services, National Cartography & Geospatial Center

APPENDIX B

Attributes and Descriptions of SSURGO Data Layers

The following attribute descriptions were taken directly from the United States Department of Agriculture Natural Resources Conservation Service Soil Data Viewer Application.

Available water capacity (AWC) refers to the quantity of water that the soil is capable of storing for use by plants. The capacity for water storage is given in centimeters of water per centimeter of soil for each soil layer. The capacity varies, depending on soil properties that affect retention of water. The most important properties are the content of organic matter, soil texture, bulk density, and soil structure, with corrections for salinity and rock fragments. Available water capacity is an important factor in the choice of plants or crops to be grown and in the design and management of irrigation systems. It is not an estimate of the quantity of water actually available to plants at any given time.

Available water supply (AWS) is computed as AWC times the thickness of the soil. For example, if AWC is 0.15 cm/cm, the available water supply for 25 centimeters of soil would be 0.15 x 25, or 3.75 centimeters of water.

For each soil layer, AWC is recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Available water supply (AWS) is the total volume of water (in centimeters) that should be available to plants when the soil, inclusive of rock fragments, is at field capacity. It is commonly estimated as the amount of water held between field capacity and the wilting point, with corrections for salinity, rock fragments, and rooting depth. AWS is reported as a single value (in centimeters) of water for the specified depth of the soil. AWS is calculated as the available water capacity times the thickness of each soil horizon to a specified depth.

For each soil layer, available water capacity, used in the computation of AWS, is recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For the derivation of AWS, only the representative value for available water capacity is used.

The available water supply for each map unit component is computed as described above and then aggregated to a single value for the map unit by the process described below.

A map unit typically consists of one or more "components." A component is either some type of soil or some nonsoil entity, e.g., rock outcrop. For the attribute being aggregated (e.g., available water supply), the first step of the aggregation process is to derive one attribute value for each of a map unit's components. From this set of component attributes, the next step of the process is to derive a single value that represents the map unit as a whole. Once a single value for each map unit is derived, a thematic map for the map units can be generated. Aggregation is needed because map units rather than components are delineated on the soil maps.

The composition of each component in a map unit is recorded as a percentage. A composition of 60 indicates that the component typically makes up approximately 60 percent of the map unit.

For the available water supply, when a weighted average of all component values is computed, percent composition is the weighting factor.

Bulk density, 15 bar, is the ovendry weight of the soil material less than 2 millimeters in size per unit volume of soil at water tension of 15 bars, expressed in grams per cubic centimeter. Bulk density, 15 bar, is necessary for resource assessment models, such as soil hydrology, water budgets, leaching,¼ and nutrient-pesticide loading.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

For each soil horizon or layer in the underlying database, this attribute is actually recorded as three separate values. A low value and a high value indicate the range of this attribute for the corresponding component. A "representative" value indicates the expected value of this attribute for the corresponding component. For this soil property, only the representative value is used.

Bulk density, one-third bar, is the ovendry weight of the soil material less than 2 millimeters in size per unit volume of soil at water tension of 1/3 bar, expressed in grams per cubic centimeter. Bulk density data are used to compute linear extensibility, shrink-swell potential, available water capacity, total pore space, and other soil properties. The moist bulk density of a soil indicates the pore space available for water and roots. Depending on soil texture, a bulk density of more than 1.4 can restrict water storage and root penetration. Moist bulk density is influenced by texture, kind of clay, content of organic matter, and soil structure.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Calcium carbonate equivalent is the percent of carbonates, by weight, in the fraction of the soil less than 2 millimeters in size. The availability of plant nutrients is influenced by the amount of carbonates in the soil. Incorporating nitrogen fertilizer into calcareous soils helps to prevent nitrite accumulation and ammonium-N volatilization.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Cation-exchange capacity (CEC-7) is the total amount of extractable bases that can be held by the soil, expressed in terms of milliequivalents per 100 grams of soil at neutrality (pH 7.0) or at some other stated pH value. Soils having a low cation-exchange capacity hold fewer cations and may require more frequent applications of fertilizer than soils having a high cation-exchange capacity. The ability to retain cations reduces the hazard of ground-water pollution.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Clay as a soil separate consists of mineral soil particles that are less than 0.002 millimeter in diameter. The estimated clay content of each soil layer is given as a percentage, by weight, of the soil material that is less than 2 millimeters in diameter. The amount and kind of clay affect the fertility and physical condition of the soil and the ability of the soil to adsorb cations and to retain moisture. They influence shrink-swell potential, saturated hydraulic conductivity (Ksat), plasticity, the ease of soil dispersion, and other soil properties. The amount and kind of clay in a soil also affect tillage and earth-moving operations.

Most of the material is in one of three groups of clay minerals or a mixture of these clay minerals. The groups are kaolinite, smectite, and hydrous mica, the best known member of which is illite.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Depth to Soil Restriction. A "restrictive layer" is a nearly continuous layer that has one or more physical, chemical, or thermal properties that significantly impede the movement of water and air through the soil or that restrict roots or otherwise provide an unfavorable root environment. Examples are bedrock, cemented layers, dense layers, and frozen layers.

This theme presents the depth to the user selected type of restrictive layer as described in for each map unit. If no restrictive layer is described in a map unit, it is represented by the "> 200" depth class.

This attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Depth to Water Table, December and July. "Water table" refers to a saturated zone in the soil. It occurs during specified months. Estimates of the upper limit are based mainly on observations of the water table at selected sites and on evidence of a saturated zone, namely grayish colors (redoximorphic features) in the soil. A saturated zone that lasts for less than a month is not considered a water table.

This attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

"**Drainage class** (natural)" refers to the frequency and duration of wet periods under conditions similar to those under which the soil formed. Alterations of the water regime by human activities, either through drainage or irrigation, are not a consideration unless they have significantly changed the morphology of the soil. Seven classes of natural soil drainage are recognized-excessively drained, somewhat excessively drained, well drained, moderately well drained, somewhat poorly drained, poorly drained, and very poorly drained. These classes are defined in the "Soil Survey Manual."

Effective cation-exchange capacity refers to the sum of extractable bases plus aluminum expressed in terms of milliequivalents per 100 grams of soil. It is determined for soils that have pH of less than 5.5. Soils having a low cation-exchange capacity (CEC) hold fewer cations and may require more frequent applications of fertilizer than soils having a high cation-exchange capacity. The ability to retain cations reduces the hazard of ground-water pollution. Effective CEC is a measure of CEC that is particularly useful in areas where the ion-exchange capacity of the soil is largely a result of variable charge components, such as allophane, kaolinite, hydrous iron and aluminum oxides, and organic matter, which result in a CEC that is not a fixed number but a function of pH.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Electrical conductivity (EC) is the electrolytic conductivity of an extract from saturated soil paste, expressed as millimhos per centimeter at 25 degrees C. Electrical conductivity is a measure of the concentration of water-soluble salts in soils. It is used to indicate saline soils. High concentrations of neutral salts, such as sodium chloride and sodium sulfate, may interfere with the absorption of water by plants because the osmotic pressure in the soil solution is nearly as high as or higher than that in the plant cells.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Farmland classification identifies map units as prime farmland, farmland of statewide importance, farmland of local importance, or unique farmland. It identifies the location and extent of the soils that are best suited to food, feed, fiber, forage, and oilseed crops. NRCS policy and procedures on prime and unique farmlands are published in the "Federal Register," Vol. 43, No. 21, January 31, 1978.

Flooding Frequency Class. Flooding is the temporary inundation of an area caused by overflowing streams, by runoff from adjacent slopes, or by tides. Water standing for short periods after rainfall or snowmelt is not considered flooding, and water standing in swamps and marshes is considered ponding rather than flooding.

Frequency is expressed as none, very rare, rare, occasional, frequent, and very frequent.

"None" means that flooding is not probable. The chance of flooding is nearly 0 percent in any year. Flooding occurs less than once in 500 years.

"Very rare" means that flooding is very unlikely but possible under extremely unusual weather conditions. The chance of flooding is less than 1 percent in any year.

"Rare" means that flooding is unlikely but possible under unusual weather conditions. The chance of flooding is 1 to 5 percent in any year.

"Occasional" means that flooding occurs infrequently under normal weather conditions. The chance of flooding is 5 to 50 percent in any year.
"Frequent" means that flooding is likely to occur often under normal weather conditions. The chance of flooding is more than 50 percent in any year but is less than 50 percent in all months in any year.

"Very frequent" means that flooding is likely to occur very often under normal weather conditions. The chance of flooding is more than 50 percent in all months of any year.
Frost Action. Potential for frost action is the likelihood of upward or lateral expansion of the soil caused by the formation of segregated ice lenses (frost heave) and the subsequent collapse of the soil and loss of strength on thawing.

Frost action occurs when moisture moves into the freezing zone of the soil. Temperature, texture, density, saturated hydraulic conductivity (Ksat), content of organic matter, and depth to the water table are the most important factors considered in evaluating the potential for frost action. It is assumed that the soil is not insulated by vegetation or snow and is not artificially drained. Silty and highly structured, clayey soils that have a high water table in winter are the most susceptible to frost action. Well drained, very gravelly, or very sandy soils are the least susceptible. Frost heave and low soil strength during thawing cause damage to pavements and other rigid structures.

Frost-Free Days. The term "frost-free days" refers to the expected number of days between the last freezing temperature (0 degrees Celsius) in spring (January-July) and the first freezing temperature in fall (August-December). The number of days is based on the probability that the values for the standard "normal" period of 1961 to 1990 will be exceeded in 5 years out of 10.

This attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this attribute, only the representative value is used.

Hydric Rating by Map Unit. This rating provides an indication of the proportion of the map unit that meets the criteria for hydric soils. Map units that are dominantly made up of hydric soils may have small areas, or inclusions, of nonhydric soils in the higher positions on the landform, and map units dominantly made up of nonhydric soils may have inclusions of hydric soils in the lower positions on the landform.

Hydric soils are defined by the National Technical Committee for Hydric Soils (NTCHS) as soils that formed under conditions of saturation, flooding, or ponding long enough during the growing season to develop anaerobic conditions in the upper part (Federal Register, 1994). Under natural conditions, these soils are either saturated or inundated long enough during the growing season to support the growth and reproduction of hydrophytic vegetation.

The NTCHS definition identifies general soil properties that are associated with wetness. In order to determine whether a specific soil is a hydric soil or nonhydric soil, however, more specific information, such as information about the depth and duration of the water table, is needed. Thus, criteria that identify those estimated soil properties unique to hydric soils have been established (Federal Register, 2002). These criteria are used to identify map unit components that normally are associated with wetlands. The criteria used are selected estimated soil properties that are described in "Soil Taxonomy" (Soil Survey Staff, 1999) and "Keys to Soil Taxonomy" (Soil Survey Staff, 2006) and in the "Soil Survey Manual" (Soil Survey Division Staff, 1993).

If soils are wet enough for a long enough period of time to be considered hydric, they should exhibit certain properties that can be easily observed in the field. These visible properties are indicators of hydric soils. The indicators used to make onsite determinations of hydric soils are specified in "Field Indicators of Hydric Soils in the United States" (Hurt and Vasilas, 2006).

References:

Federal Register. July 13, 1994. Changes in hydric soils of the United States.

Federal Register. September 18, 2002. Hydric soils of the United States.

Hurt, G.W., and L.M. Vasilas, editors. Version 6.0, 2006. Field indicators of hydric soils in the United States.

Soil Survey Division Staff. 1993. Soil survey manual. Soil Conservation Service. U.S. Department of Agriculture Handbook 18.

Soil Survey Staff. 1999. Soil taxonomy: A basic system of soil classification for making and interpreting soil surveys. 2nd edition. Natural Resources Conservation Service. U.S. Department of Agriculture Handbook 436.

Soil Survey Staff. 2006. Keys to soil taxonomy. 10th edition. U.S. Department of Agriculture, Natural Resources Conservation Service.

Hydrologic soil groups are based on estimates of runoff potential. Soils are assigned to one of four groups according to the rate of water infiltration when the soils are not protected by vegetation, are thoroughly wet, and receive precipitation from long-duration storms.

The soils in the United States are assigned to four groups (A, B, C, and D) and three dual classes (A/D, B/D, and C/D). The groups are defined as follows:

Group A. Soils having a high infiltration rate (low runoff potential) when thoroughly wet. These consist mainly of deep, well drained to excessively drained sands or gravelly sands. These soils have a high rate of water transmission.

Group B. Soils having a moderate infiltration rate when thoroughly wet. These consist chiefly of moderately deep or deep, moderately well drained or well drained soils that have moderately fine texture to moderately coarse texture. These soils have a moderate rate of water transmission.

Group C. Soils having a slow infiltration rate when thoroughly wet. These consist chiefly of soils having a layer that impedes the downward movement of water or soils of moderately fine texture or fine texture. These soils have a slow rate of water transmission.

Group D. Soils having a very slow infiltration rate (high runoff potential) when thoroughly wet. These consist chiefly of clays that have a high shrink-swell potential, soils that have a high water table, soils that have a claypan or clay layer at or near the surface, and soils that are shallow over nearly impervious material. These soils have a very slow rate of water transmission.

If a soil is assigned to a dual hydrologic group (A/D, B/D, or C/D), the first letter is for drained areas and the second is for undrained areas. Only the soils that in their natural condition are in group D are assigned to dual classes.

K Factor, Rock Free and Whole Soil. Erosion factor K indicates the susceptibility of a soil to sheet and rill erosion by water. Factor K is one of six factors used in the Universal Soil Loss Equation (USLE) and the Revised Universal Soil Loss Equation (RUSLE) to predict the average annual rate of soil loss by sheet and rill erosion in tons per acre per year. The estimates are based primarily on percentage of silt, sand, and organic matter and on soil structure and saturated hydraulic conductivity (Ksat). Values of K range from 0.02 to 0.69. Other factors being equal, the higher the value, the more susceptible the soil is to sheet and rill erosion by water.

"Erosion factor Kf (rock free)" indicates the erodibility of the fine-earth fraction, or the material less than 2 millimeters in size.

"Erosion factor Kw (whole soil)" indicates the erodibility of the whole soil. The estimates are modified by the presence of rock fragments.

Liquid limit (LL) is one of the standard Atterberg limits used to indicate the plasticity characteristics of a soil. It is the water content, on a percent by weight basis, of the soil (passing #40 sieve) at which the soil changes from a plastic to a liquid state. Generally, the amount of clay- and silt-size particles, the organic matter content, and the type of minerals determine the liquid limit. Soils that have a high liquid limit have the capacity to hold a lot of water while maintaining a plastic or semisolid state.

Liquid limit is used in classifying soils in the Unified and AASHTO classification systems.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Organic matter is the plant and animal residue in the soil at various stages of decomposition. The estimated content of organic matter is expressed as a percentage, by weight, of the soil material that is less than 2 millimeters in diameter.

The content of organic matter in a soil can be maintained by returning crop residue to the soil. Organic matter has a positive effect on available water capacity, water infiltration, soil organism activity, and tilth. It is a source of nitrogen and other nutrients for crops and soil organisms. An irregular distribution of organic carbon with depth may indicate different episodes of soil deposition or soil formation. Soils that are very high in organic matter have poor engineering properties and subside upon drying.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

pH (1 to 1 water) Soil reaction is a measure of acidity or alkalinity. It is important in selecting crops and other plants, in evaluating soil amendments for fertility and stabilization, and in determining the risk of corrosion. In general, soils that are either highly alkaline or highly acid are likely to be very corrosive to steel. The most common soil laboratory measurement of pH is the 1:1 water method. A crushed soil sample is mixed with an equal amount of water, and a measurement is made of the suspension.
For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this

attribute for the component. For this soil property, only the representative value is used.

Plasticity index (PI) is one of the standard Atterberg limits used to indicate the plasticity characteristics of a soil. It is defined as the numerical difference between the liquid limit and plastic limit of the soil. It is the range of water content in which a soil exhibits the characteristics of a plastic solid.

The plastic limit is the water content that corresponds to an arbitrary limit between the plastic and semisolid states of a soil. The liquid limit is the water content, on a percent by weight basis, of the soil (passing #40 sieve) at which the soil changes from a plastic to a liquid state.

Soils that have a high plasticity index have a wide range of moisture content in which the soil performs as a plastic material. Highly and moderately plastic clays have large PI values. Plasticity index is used in classifying soils in the Unified and AASHTO classification systems.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Ponding Frequency Class. Ponding is standing water in a closed depression. The water is removed only by deep percolation, transpiration, or evaporation or by a combination of these processes. Ponding frequency classes are based on the number of times that ponding occurs over a given period. Frequency is expressed as none, rare, occasional, and frequent.

"None" means that ponding is not probable. The chance of ponding is nearly 0 percent in any year.

"Rare" means that ponding is unlikely but possible under unusual weather conditions. The chance of ponding is nearly 0 percent to 5 percent in any year.

"Occasional" means that ponding occurs, on the average, once or less in 2 years. The chance of ponding is 5 to 50 percent in any year.
"Frequent" means that ponding occurs, on the average, more than once in 2 years. The chance of ponding is more than 50 percent in any year.

Sand as a soil separate consists of mineral soil particles that are 0.05 millimeter to 2 millimeters in diameter. In the database, the estimated sand content of each soil layer is given as a percentage, by weight, of the soil material that is less than 2 millimeters in diameter. The content of sand, silt, and clay affects the physical behavior of a soil. Particle size is important for engineering and agronomic interpretations, for determination of soil hydrologic qualities, and for soil classification.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Saturated hydraulic conductivity (Ksat) refers to the ease with which pores in a saturated soil transmit water. The estimates are expressed in terms of micrometers per second. They are based on soil characteristics observed in the field, particularly structure, porosity, and texture. Saturated hydraulic conductivity is considered in the design of soil drainage systems and septic tank absorption fields.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

The numeric Ksat values have been grouped according to standard Ksat class limits. The classes are:

Very low: 0.00 to 0.01
Low: 0.01 to 0.1
Moderately low: 0.1 to 1.0
Moderately high: 1 to 10
High: 10 to 100
Very high: 100 to 705

Silt as a soil separate consists of mineral soil particles that are 0.002 to 0.05 millimeter in diameter. In the database, the estimated silt content of each soil layer is given as a percentage, by weight, of the soil material that is less than 2 millimeters in diameter.

The content of sand, silt, and clay affects the physical behavior of a soil. Particle size is important for engineering and agronomic interpretations, for determination of soil hydrologic qualities, and for soil classification

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Slope gradient is the difference in elevation between two points, expressed as a percentage of the distance between those points.

The slope gradient is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Sodium adsorption ratio (SAR) is a measure of the amount of sodium (Na) relative to calcium (Ca) and magnesium (Mg) in the water extract from saturated soil paste. It is the ratio of the Na concentration divided by the square root of one-half of the Ca + Mg concentration. Soils that have SAR values of 13 or more may be characterized by an increased dispersion of organic matter and clay particles, reduced saturated hydraulic conductivity (Ksat) and aeration, and a general degradation of soil structure.

For each soil layer, this attribute is actually recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Surface Texture. This displays the representative texture class and modifier of the surface horizon.

Texture is given in the standard terms used by the U.S. Department of Agriculture. These terms are defined according to percentages of sand, silt, and clay in the fraction of the soil that is less than 2 millimeters in diameter. "Loam," for example, is soil that is 7 to 27 percent clay, 28 to 50 percent silt, and less than 52 percent sand. If the content of particles coarser than sand is 15 percent or more, an appropriate modifier is added, for example, "gravelly."

Water content, 15 bar, is the amount of soil water retained at a tension of 15 bars, expressed as a percentage of the oven-dry weight of soil material that is less than 2 mm in diameter. Water retained at 15 bars is significant in the determination of soil water-retention difference, which is used as the initial estimation of available water capacity for some soils. Water retained at 15 bars is an estimation of the wilting point.

Water content varies between soil types, depending on soil properties that affect retention of water. The most important properties are the content of organic matter, soil texture, bulk density, and soil structure.

For each soil layer, water content is recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.

Water content, one-third bar, is the amount of soil water retained at a tension of 1/3 bar, expressed as a percentage of the oven-dry weight of soil material that is less than 2 mm in diameter. Water retained at 1/3 bar is significant in the determination of soil water-retention difference, which is used as the initial estimation of available water capacity for some soils. Water retained at 1/3 bar is the value commonly used to estimate the content of water at field capacity for most soils.

Water content varies between soil types, depending on soil properties that affect retention of water. The most important properties are the content of organic matter, soil texture, bulk density, and soil structure.

For each soil layer, water content is recorded as three separate values in the database. A low value and a high value indicate the range of this attribute for the soil component. A "representative" value indicates the expected value of this attribute for the component. For this soil property, only the representative value is used.